物联网安全理论与技术

杨奎武　郑康锋　张冬梅　郭渊博　编著

電子工業出版社
Publishing House of Electronics Industry
北京·BEIJING

VISIT…

LANZAROTE Caliente.com

内 容 简 介

本书以无线传感器网络和 RFID 系统为典型代表，从物联网感知层、网络层、应用层安全三方面对物联网面临的安全问题及最新的物联网安全理论和技术进行了深入介绍。尤其是在感知层安全方面，本书更是从物理安全、认证机制、密钥管理、传输安全、协议安全、入侵检测、系统安全七个方面，以攻防相结合、理论与实践相结合的方法重点阐述了物联网面临的安全问题及解决方案。同时，本书还给出了诸如 PUF 技术、物理层秘钥生成、移动目标防御等近年最新的学术研究成果。

本书内容丰富、涵盖面广、系统性强，非常适合作为物联网、通信、信息安全等相关专业的高年级本科生、研究生的教材及参考书，也非常适合作为物联网及信息安全领域教师和工程技术人员的参考用书。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

物联网安全理论与技术/杨奎武等编著. —北京：电子工业出版社，2017.1
ISBN 978-7-121-30407-1

Ⅰ. ① 物…　Ⅱ. ① 杨…　Ⅲ. ① 互联网络 – 应用 – 安全技术　② 智能技术 – 应用 – 安全技术
Ⅳ. ① TP393.4　② TP18

中国版本图书馆 CIP 数据核字（2016）第 279661 号

策划编辑：曲　昕
责任编辑：谭丽莎
印　　刷：
装　　订：
出版发行：电子工业出版社
　　　　　北京市海淀区万寿路 173 信箱　邮编 100036
开　　本：787×1 092　1/16　印张：17　字数：435.2 千字
版　　次：2017 年 1 月第 1 版
印　　次：2017 年 1 月第 1 次印刷
定　　价：49.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010)88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：quxin@phei.com.cn。

前　言

物联网的概念从提出到现在已经有近 20 年了，虽然刚刚提出来的时候很多人都觉得物联网很遥远，但是随着近年来科学技术的快速发展，我们现在已经能够切身体会到物联网的无处不在了。在物联网进一步普适化、泛在化的道路上，恐怕目前技术领域的唯一绊脚石就是安全问题了，这也自然成为物联网领域研究的热点。

为适应物联网安全工作的发展需要，满足广大物联网技术领域学生、教师及工程技术人员的学习工作需求，我们结合多年课题研究的成果编写了本书，目的就是为广大读者提供参考，进一步推动物联网的广泛应用。

本书从物联网安全基础、物联网感知安全、物联网网络安全、物联网应用安全四个层面对物联网安全问题及典型的物联网安全技术进行了总结和分析。全书共 12 章，第 1、2 章主要介绍了物联网的基本概念、结构，物联网安全面临的挑战及安全框架，以及物联网安全技术的密码学基础知识。第 3 ~ 9 章分别从物理安全、认证机制、密钥管理、数据传输安全、MAC 层协议安全、感知层入侵检测和嵌入式系统安全 7 个方面介绍了物联网感知层安全的关键技术，这部分内容也是本书的主体和重点，其中包含了很多作者最新的研究成果。第 10 章介绍物联网接入网络安全技术，主要包括无线局域网、WiMAX 和移动通信接入网络的安全问题。第 11 章主要从被动防御和主动防御两个方面介绍了物联网核心网络的典型安全技术。第 12 章重点从物联网云安全的角度给出了物联网应用相关安全技术介绍。

本书由杨奎武主笔并统稿，由杨奎武、郑康锋、张冬梅和郭渊博共同编撰完成，编写过程中得到了信息工程大学和北京邮电大学的大力支持，陈越教授、贾洪勇讲师、武斌讲师等很多教师都对本书提出了宝贵的意见和建议；傅蓉蓉、肖倩、张紫楠、隋雷、姜文博、郑巧琼、肖欢、张之则等研究生协助完成了部分文档的整理工作；电子工业出版社的曲昕编辑为本书的出版付出了辛勤的劳动，在此向大家表示由衷的感谢。另外，本书很多内容参考了互联网上的资源及其他相关著作，在此对资源的分享者和其他著作作者一并表示感谢。

由于物联网安全技术发展迅速，新技术和新标准不断涌现，加之作者水平有限，编写时间仓促，本书难免存在错误和不足之处，敬请各位专家和读者批评指正。

编著者

2017 年 1 月

目　　录

第 1 章　绪论 …… 1
1.1　概述 …… 1
1.1.1　物联网的定义 …… 1
1.1.2　物联网体系结构 …… 1
1.1.3　物联网的主要特点及应用领域 …… 3
1.1.4　物联网发展现状 …… 5
1.1.5　物联网关键技术 …… 6
1.2　物联网安全 …… 10
1.2.1　物联网安全特点及面临的安全挑战 …… 10
1.2.2　物联网面临的安全威胁与安全目标 …… 12
1.2.3　物联网安全技术框架 …… 17
参考文献 …… 19
第 2 章　密码技术基础 …… 20
2.1　密码学概述 …… 20
2.1.1　密码体制 …… 20
2.1.2　密码分类 …… 21
2.2　分组密码 …… 22
2.2.1　DES …… 22
2.2.2　AES …… 24
2.3　公钥密码体制 …… 25
2.3.1　RSA …… 25
2.3.2　ElGamal 和 ECC …… 26
2.3.3　公钥密码体制应用 …… 26
2.4　认证与数字签名 …… 27
2.4.1　Hash 函数 …… 27
2.4.2　报文认证 …… 28
2.4.3　数字签名 …… 30
2.5　密钥管理与分发 …… 31
参考文献 …… 33
第 3 章　感知层物理安全技术 …… 34

3.1 RFID 标签物理层安全威胁及防护技术 …… 34
3.1.1 RFID 标签的破解及复制 …… 34
3.1.2 RFID 标签的物理安全防护技术 …… 36
3.2 传感器网络节点的物理安全威胁及其防御技术 …… 39
3.2.1 节点破坏攻击及其防御 …… 39
3.2.2 节点泄露攻击及其防御 …… 40
3.2.3 传感器节点安全设计 …… 41
3.3 物理不可克隆函数（PUF）技术 …… 42
3.3.1 PUF 概述 …… 42
3.3.2 PUF 基本原理及其数学模型 …… 43
3.3.3 PUF 分类及实现 …… 44
3.3.4 PUF 属性 …… 47
3.3.5 PUF 研究及应用现状 …… 50
参考文献 …… 53
第 4 章　感知层认证技术 …… 56
4.1 感知层认证技术概述 …… 56
4.1.1 RFID 认证技术 …… 57
4.1.2 无线传感器网络认证技术 …… 57
4.2 RFID 认证机制 …… 58
4.2.1 基于 Hash 函数的认证机制 …… 58
4.2.2 RFID 分布式询问–应答认证机制 …… 61
4.2.3 RFID 轻量级安全认证 …… 62
4.2.4 一种基于 PUF 的 RFID 认证协议 …… 63
4.3 传感器网络认证技术 …… 65
4.3.1 SNEP 网络安全加密协议 …… 65
4.3.2 uTESLA 广播消息认证协议 …… 68
4.3.3 基于身份标识加密的身份认证 …… 71
4.3.4 基于 PUF 的延迟容忍传感器网络节点身份认证机制 …… 72
参考文献 …… 76
第 5 章　感知层密钥管理技术 …… 78
5.1 感知层密钥管理技术概述 …… 78
5.1.1 RFID 密钥管理技术 …… 78
5.1.2 传感器网络密钥管理技术 …… 78
5.2 基于 HB 协议族的 RFID 密钥协商及管理技术 …… 81
5.2.1 LPN 问题概述 …… 81
5.2.2 HB 协议 …… 81
5.2.3 HB + 协议 …… 82
5.2.4 HB ++ 协议 …… 82

5.3　传感器网络密钥分配及管理技术 …… 83
5.3.1　预共享密钥机制 …… 83
5.3.2　随机密钥分配机制 …… 84
5.3.3　分簇传感器网络的密钥管理机制 …… 88
5.3.4　基于 PUF 的 DTMSN 密钥管理机制 …… 92
5.4　基于物理层信道特征的密钥生成技术 …… 96
5.4.1　物理层安全 …… 96
5.4.2　基于信道特征的密钥生成 …… 99
5.4.3　一种无线物理层密钥生成机制 …… 102
参考文献 …… 103
第 6 章　感知层数据安全传输技术 …… 106
6.1　RFID 系统安全通信技术 …… 106
6.1.1　RFID 差错控制技术 …… 106
6.1.2　RFID 数据传输防碰撞技术 …… 109
6.2　传感器网络安全路由技术 …… 113
6.2.1　无线传感器网络路由协议概述 …… 113
6.2.2　传感器网络信息协商路由协议（SPINS） …… 116
6.2.3　INSENSE 入侵容忍路由协议 …… 118
6.2.4　协作式安全路由协议 …… 119
6.3　网络编码技术在数据传输中的应用 …… 120
6.3.1　网络编码的基本原理及分类 …… 121
6.3.2　随机网络编码技术 …… 122
6.3.3　COPE：一种实际的编码路由协议 …… 124
6.3.4　一种基于网络编码的延迟容忍移动传感器网络广播数据传输机制 …… 127
参考文献 …… 131
第 7 章　感知层 MAC 协议安全 …… 133
7.1　无线传感器网络 802.15.4 MAC 层协议 …… 133
7.1.1　IEEE 802.15.4 标准 …… 133
7.1.2　IEEE 802.15.4 网络协议栈 …… 134
7.1.3　IEEE 802.15.4 MAC 帧格式 …… 134
7.2　IEEE 802.15.4 协议安全分析 …… 137
7.2.1　信标广播机制及安全分析 …… 137
7.2.2　GTS 管理机制及安全分析 …… 138
7.3　无线局域网概述 …… 140
7.3.1　无线局域网的基本构成 …… 140
7.3.2　无线局域网网络结构 …… 140
7.3.3　IEEE 802.11 相关标准 …… 141
7.3.4　IEEE 802.11 协议体系 …… 143

7.4　无线局域网 MAC 层接入认证协议 …… 143
7.4.1　WEP 身份认证协议 …… 143
7.4.2　WPA/WPA2 - PSK 认证机制 …… 144
7.4.3　IEEE 802.1x/EAP 认证机制 …… 146
7.5　无线局域网 MAC 层协议安全分析 …… 151
7.5.1　WEP 中的安全隐患 …… 151
7.5.2　WPA/WPA - PSK 认证协议安全分析 …… 151
7.5.3　IEEE 802.1x/EAP 认证协议安全分析 …… 152
参考文献 …… 153
第 8 章　感知层入侵检测技术 …… 155
8.1　物联网入侵检测技术概述 …… 155
8.1.1　物联网入侵检测概述 …… 155
8.1.2　常见的物联网入侵检测技术 …… 155
8.2　通用型入侵检测算法 …… 157
8.2.1　基于分簇的入侵检测算法 …… 157
8.2.2　基于博弈论的入侵检测算法 …… 160
8.2.3　基于模糊理论的阻塞攻击入侵检测算法 …… 160
8.2.4　基于人工免疫的入侵检测技术 …… 164
参考文献 …… 168
第 9 章　感知层嵌入式系统安全 …… 171
9.1　平台安全——可信计算技术 …… 171
9.1.1　可信计算技术概述 …… 171
9.1.2　TCG 可信计算平台体系结构及特征 …… 172
9.1.3　TPM 可信平台模块 …… 176
9.2　平台安全——TrustZone 技术 …… 177
9.2.1　TrustZone 技术概述 …… 177
9.2.2　TrustZone 硬件架构 …… 178
9.2.3　TrustZone 软件架构 …… 180
9.3　TinyOS 操作系统及其安全技术 …… 182
9.3.1　TinyOS 操作系统概述 …… 182
9.3.2　TinySEC 传感器网络安全体系结构 …… 186
参考文献 …… 189
第 10 章　感知层无线接入网络安全技术 …… 191
10.1　无线局域网安全保密体系结构及实现 …… 191
10.1.1　无线局域网安全目标 …… 191
10.1.2　主要安全威胁 …… 193
10.1.3　无线局域网安全需求 …… 195
10.1.4　需要的安全措施 …… 197

10.1.5　安全无线局域网的基本结构和实现方案 …… 197
10.2　WiMAX 安全技术 …… 201
10.2.1　WiMAX 网络概述 …… 201
10.2.2　WiMAX 安全体系架构 …… 205
10.2.3　IEEE 802.16m 安全机制 …… 206
10.3　3G 和 LTE 安全技术 …… 212
10.3.1　3G 移动通信网络及安全威胁 …… 212
10.3.2　3GPP 安全增强技术 …… 213
10.3.3　LTE/SAE（4G）安全技术 …… 215
参考文献 …… 221
第 11 章　物联网核心网络安全技术 …… 222
11.1　被动防御——计算机病毒检测技术 …… 222
11.1.1　计算机病毒 …… 222
11.1.2　计算机病毒的特点及分类 …… 223
11.1.3　计算机病毒检测技术 …… 225
11.2　被动防御——防火墙技术 …… 227
11.2.1　防火墙的概念 …… 227
11.2.2　防火墙的分类 …… 227
11.2.3　防火墙的配置 …… 228
11.3　主动防御——入侵检测技术 …… 230
11.3.1　IDS 的标准结构 …… 231
11.3.2　IDS 的分类 …… 231
11.4　主动防御——网络态势感知技术 …… 234
11.4.1　网络态势感知研究框架 …… 235
11.4.2　网络态势感知模型 …… 236
11.4.3　网络态势知识表示 …… 237
11.4.4　评估方法分类 …… 237
11.5　主动防御——移动目标防御技术 …… 238
11.5.1　移动目标、移动目标防御及拟态安全防御 …… 239
11.5.2　移动目标防御技术的最新进展 …… 240
11.5.3　移动目标防御机制 …… 241
参考文献 …… 243
第 12 章　物联网应用层云安全技术 …… 246
12.1　云计算简介 …… 246
12.1.1　云计算的概念 …… 246
12.1.2　云计算的特点 …… 246
12.1.3　云计算的分类 …… 247
12.2　物联网与云计算的融合 …… 248

12.2.1　与云计算相融合是发展必然 …… 248
12.2.2　基于云计算的物联网系统 …… 249
12.2.3　云计算与物联网的融合模式 …… 250
12.3　云计算安全问题 …… 251
12.3.1　IaaS 安全问题 …… 251
12.3.2　PaaS 安全问题 …… 252
12.3.3　SaaS 安全问题 …… 253
12.3.4　其他安全问题 …… 254
12.4　云安全关键技术 …… 255
12.5　基于云计算的物联网信息安全服务体系 …… 261
参考文献 …… 262

第1章 绪　　论

随着物联网技术研究和应用的不断发展，物联网安全面临的问题越来越突出，成为制约网络发展的重要瓶颈，这也使得物联网安全及隐私保护技术日益成为国内外学者研究的焦点。相对于传统的计算机网络安全技术，物联网安全及隐私保护技术研究涉及的内容更加广泛，也更具复杂性。本章在对物联网的基本概念、体系结构及关键技术等内容介绍的基础上，从总体上讲述了物联网面临的安全及挑战，给出了物联网的安全目标及当前物联网安全技术研究的热点。

1.1 概述

1.1.1 物联网的定义

物联网（Internet of Things，IoT）这一概念，是由麻省理工学院自动识别实验室于1999年在研究RFID时提及并引起关注的，并于2005年11月国际电信联盟（ITU）发布的《ITU互联网报告2005：物联网》中正式提出并进行扩展。针对物联网的定义，目前国际上并没有一致认同的准确和权威的定义，并且随着技术的进步，物联网的定义及其所涉及的内涵和外延也都在不断发生变化。

目前在国内被最普遍引用的物联网定义是：通过射频识别（RFID）、红外感应器、全球定位系统、激光扫描器等信息传感设备，按约定的协议，把任何物品与互联网连接起来，进行信息交换和通信，以实现智能化识别、定位、跟踪、监控和管理的一种网络。它是在互联网基础上延伸和扩展的网络。

而“全球RFID运作及标准化协调支持行动（CASAGRAS）”项目给出的物联网的定义如下：IoT is a global network infrastructure, linking physical and virtual objects through the exploitation of data capture and communication capabilities. This infrastructure includes existing and evolving Internet and network developments. It will offer specific object - identification, sensor and connection capability as the basis for the development of independent federated services and applications. These will be characterized by a high degree of autonomous data capture, event transfer, network connectivity and interoperability.

虽然国际上对物联网的定义并没有统一认识，但通俗地讲，物联网就是一个通过信息技术将各种物体连接成网络，使物体变得更加智能化，从而实现人与物、物与物之间的通信的网络。物联网对其所连接的物体主要有三点要求：一是物联网中每一个物体都可寻址；二是每一个物体均可以通信；三是每一个物体均可控制。

1.1.2 物联网体系结构

物联网的体系结构目前较为公认的是三层体系结构，即物联网从下到上分为三个层次，依次是：感知层、网络层（包括接入网络）和应用层（包括信息处理、云计算等平台），如

图 1-1 所示。

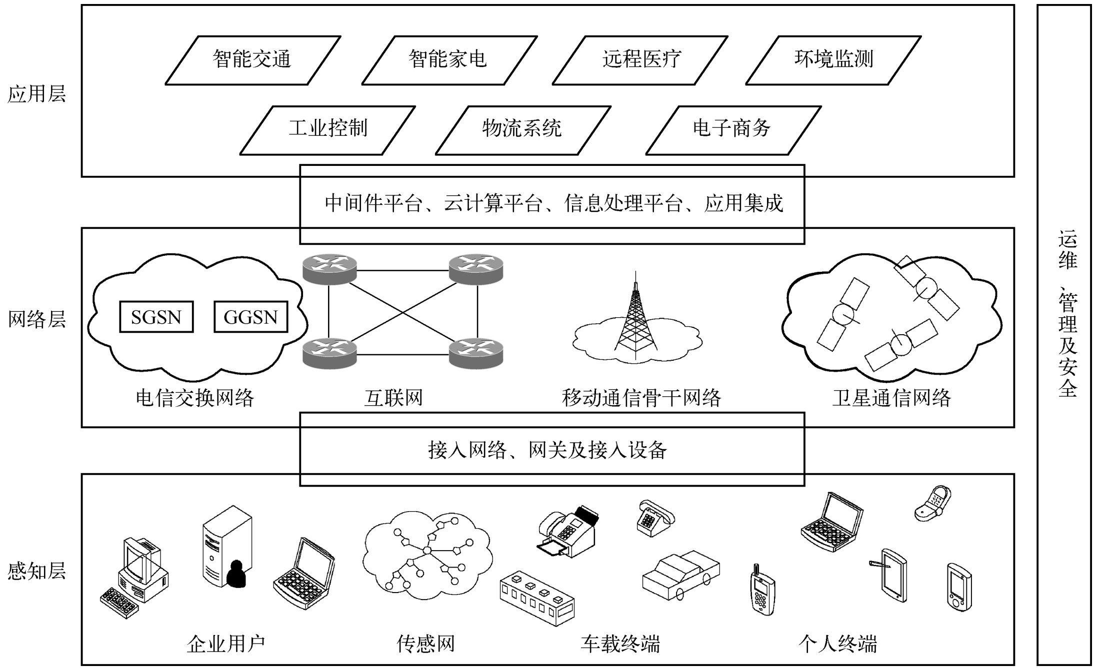


图 1-1　物联网体系结构

1. 感知层

物联网的感知层处于网络的边沿，分布最为广泛，如果将物联网比作人，则感知层就相当于人体的感知器官，主要用来完成信息的感知和采集。感知层设备多种多样，种类非常丰富，常见的有 RFID、传感器节点、红外感知设备、摄像头、智能手机、各种传感器设备，以及由这些设备组成的同构或异构的网络。根据网络应用的不同，采用的感知设备可能不同。例如智能农业系统，感知层设备主要由采集土壤温度、湿度、光照的传感器设备组成；而在物流系统中，感知层设备主要由 RFID 标签、阅读器组成；在智能交通系统中，感知层设备主要由测速仪器、GPS 定位终端、超声雷达等组成。感知层是物联网信息的来源和起点，直接与被监控的“物”相连。物联网就是通过各类传感器、RFID 及其他感知设备使得物体能够智能地表达自己，从而实现物与物之间的智能通信的。

2. 网络层

网络层是物联网的核心部分，属于网络的中枢神经系统，负责信息的传输和交互。物联网的网络层主要以互联网、移动通信网络、电信交换网络、卫星通信网络为主要组成部分，用于实现网络中各类终端设备的广泛互连，实现感知层数据的高效、可靠、安全传输，同时为各类感知层异构网络提供接入接口，实现网络层与感知层的紧密融合。电信网络、互联网技术成熟度较高，是物联网的网络层的骨干，也是物联网各类业务的主要支撑环境，但针对未来新型的物联网业务，如 RFID 标签位置查询等，现有的骨干网络结构仍需须进一步调整、优化。

感知层与网络层之间的互连互通主要通过接入网络来实现。接入网络主要有 WLAN、WiMAX、MESH 等多种网络形式，能够支持多种通信标准。感知层设备可以直接与接入网络相连，也可以通过网关与接入网络相连，接入网络再通过专用设备与骨干网络相通。

3. 应用层

物联网的应用层主要实现各类的应用服务，如智能交通、环境监测、物流管理、智能楼宇、电子支付等，不同的应用服务可能对应于网络层及感知层的不同功能。例如在智能交通应用中，用户可以通过移动通信网络或 WiMAX 网络等实时获取当前路况信息，共享娱乐资讯，利用短距离通信手段实现车与车之间的对话，同时还可以通过车内设置的各类传感器节点时刻获取车内发动机、油路、空调等系统的运行情况，并自动将车内异常情况远程上传到车辆服务中心获取技术支持。相对于互联网而言，物联网的应用类型和服务更为多样和丰富。因此物联网不仅仅是通信技术的革新，更是人类生活方式的变革。

近年来，为了进一步支撑物联网应用服务的发展，很多大公司还在网络层与应用层之间构建了专门的数据或信息处理平台，如云存储平台、云计算平台、数据分析平台等，这些平台的建立，有效降低了应用服务的成本，为应用服务的多样化提供了基础。

物联网体系结构除了以上三个层次外，还涉及一些公共技术，这些公共技术与三个层次都有关系，提供网络的管理、维护，数据的查询、挖掘，行为的分析、决策，安全的防护、检测等。

1.1.3　物联网的主要特点及应用领域

1. 物联网的主要特点

中国移动前董事长王建宙表示，物联网有三大特征：一是全面感知，即利用 RFID、传感器、二维码、GPS 等设备或标签随时随地获取物体的信息；二是可靠传递，通过电信网络、互联网的融合，将信息实时、准确地传输到目的地；三是智能处理，利用云计算、智能识别、人工智能等各种技术，对海量信息和数据进行分析、处理，实现对物体的智能化控制。

互联网是人与人之间的网络，而物联网是物与物、物与人之间的网络，因此除了全面感知、可靠传递和智能处理这三个重要特征外，相对互联网而言，物联网还有以下特点。

1）对象更广

除了人以外，相对互联网而言，物联网所涉及的对象更多，生产、生活中的物品，如服装、手表、汽车、车床、仪表等物品都可以接入网络，实现智能化通信。

2）范围更大

物联网包含互联网、移动通信网、卫星通信网等多种网络，其覆盖范围更加广泛，接入手段更加多样，终端可以静止也可以移动，终端设备可以在任何地点，通过多种接入方式实现网络的接入。

3）智能更高

通过使用大数据分析、云计算等基础平台，物联网用户可以随时随地分享平台成果，获

得强大的平台支撑，从而更为智能地处理复杂的事件。

4）能力更强

随着技术的发展，尤其是处理器技术的不断进步，物联网终端的成本不断降低、性能不断提升，能够完成的功能也日益多样。

2. 物联网的典型应用

物联网的应用前景非常广阔，遍及智能交通、路桥管理、灾害防治、卫生保健、儿童护理、环境监管、平安家居、敌情侦察、情报获取、工业控制等多个领域。物联网的出现使物品和服务都发生了质的变化，改变了人们的生活方式，为使用者提供了更高的效率。物联网最终将发展成为面向服务的网络，根据用户的需求来提供智能化的便捷服务。这一技术将会发展成为一个有着上万亿规模的高科技市场。目前典型的物联网应用如下。

1）智能环保

随着经济的发展，人们对生活质量和环境的要求越来越高。为了提高环境监测和管理水平，环保部门可以建设基于物联网的智能环保通信系统。通过建设形成一个覆盖全区的环境自动监测信息采集网络，实现对重点排污单位污染防治设施运行状态、主要污染物排放检测数据的自动传输和预警，实现重点流域水环境质量、重点城市环境空气质量自动监测数据实时传输。通过建设一个环境分析系统和一个交互式的环境监测、环境保护的动态信息发布平台向政府、公众发布环境信息，实现集环境监测的智能感知、智能处理和综合管理于一体，推进污染减排和环境保护，实现环境与人、经济乃至整个社会的协调发展，促进环境改善。

2）智能物流

RFID 技术是物联网的重要技术基础，是实现智能物流的重要手段。利用 RFID 技术，将物流中的物品贴上电子标签，利用标签阅读器便可以在物品运转的各个环境实现对物品的清点、查询和统计，节省了人力资源。同时利用电子标签，可以随时跟踪物品所处的位置、流通环节、出厂时间，也可以方便消费者对用户进行信息检索，获取商品的来源、生产日期、主要成分及其他相关信息。物联网技术能够大大增强物流中运输、保管、装卸、包装、流通加工等物流环境的功能，使物流与商流、资金流、信息流融为一体，提升生产、流通和消费的综合效益，促进物流成本的不断下降。如今，打开手机应用我们就可以知道购买的商品所处的物理阶段，随着技术的进一步发展，我们还有可能随时定位商品的位置。

3）智能交通

以往的城市交通管理基本上都是自发进行的，驾驶者根据自己的判断选择行车路线，交通提示牌、指示灯的作用非常有限。随着 GPS、RFID 等物联网技术的发展，互联网公司能够随时获得城市车辆的位置、速度等信息，从而智能地对交通情况进行分析并给出出行参考，未来这些信息与交通管理和调度机制进一步融合，就能够充分发挥道路基础设施的效能，最大化交通网络流量，并提高交通安全性，提升人们的出行体验，甚至可以降低污染排放，提升路途乐趣。

4）军事应用

美国著名军事预测学家詹姆斯·亚当斯在其所著的《下一场世界战争》中曾预言：“在

未来的战争中，计算机本身就是武器，前线无处不在。夺取作战空间控制权的不是子弹，而是计算机网络里流动的比特和字节。”物联网可以有效应用于战场情报获取、军事物资管理等领域。通过飞机抛洒传感器节点，可以实现无人区或敌占区目标信息的采集，可以感知目标区域人员、车辆的运动趋势，甚至判断出具体目标类型，为军事决策提供准确的情报来源。物联网可以应用于战争准备、战斗实施的每一个环境，能够在多种场合满足军事信息获取的实时性、准确性、全面性的需求。

1.1.4　物联网发展现状

1. 国外物联网发展现状

虽然很多资料都认为物联网的概念是 1999 年麻省理工 RFID 实验室 Kevin Ashton 教授提出的，但其实在此之前的 1995 年，微软创始人比尔·盖茨就已经在其《The Road Ahead》一书中提到物联网的概念，只不过受限于当时的技术条件并未引起大家的重视。随着技术的进步，2005 年 11 月国际电信联盟在突尼斯举行的信息社会世界峰会上发布了《ITU 互联网报告 2005：物联网》，引用了“物联网”的概念，并对其进行了扩展，提出任何时刻、任何地点的任何物体之间都可以进行互连，无处不在的“物联网”通信时代即将来临，世界上的所有物体都可以通过互联网实现信息交换。传感器网络、射频识别、嵌入式、纳米技术将得到更加广泛的应用。

IBM 公司 2008 年提出了“智慧地球”的概念，其本质是以一种更智慧的方法，利用新一代的信息通信技术来改变政府、公司和人们相互交换的方式，以提高交换的明确性、灵活性和效率。“智慧地球”在技术层面上是物联网与互联网的融合，从而使人类能够以更加精细和动态的方式管理生产和生活，形成“物联网 + 互联网 = 智慧的地球”。“智慧地球”体现了智能城市、智能家居、智能货运、智能交通、智能医疗等多个方面。这一概念得到奥巴马政府的积极响应，物联网已经上升到美国国家发展战略层面，并引起全世界的广泛关注。

2009 年 2 月，奥巴马签署生效的《2009 年美国恢复和再投资法案》（即美国的经济刺激计划）提出要在智能电网领域应用物联网，例如得克萨斯州的电网公司建立了智能的数字电网。

2009 年 7 月，日本 IT 战略本部颁布了日本新一代的信息化战略——i – Japan，让数字技术融入社会的每一个角落；并且提出到 2015 年使行政流程简单化、效率化、标准化、透明化，并推动远程医疗和远程教育的发展。

2009 年 9 月，欧盟第 7 框架下的 RFID 和物联网研究项目组发布了《物联网战略研究路线图》研究报告，认为物联网是未来 Internet 的一个组成部分，是一种动态的全球网络基础架构，它基于标准的、可互操作的通信协议，具有自我配置能力。物联网中的“物”都具有标识，拥有物理属性，使用智能接口，能够实现与信息网络的无缝结合。

2009 年 10 月，韩国通过了物联网基础设施构建规划，将物联网市场确定为新的经济增长动力。

2013 年 4 月的汉诺威工业博览会上，德国政府提出“工业 4.0”战略，其目的是为了提高德国工业的竞争力，在新一轮工业革命中占领先机。该战略已经得到德国科研机构和产业

界的广泛认同，西门子公司已经开始将这一概念引入其工业软件开发和生产控制系统。

2. 国内物联网发展现状

我国对“物联网”的发展给予了高度重视。目前，我国对物联网的研发聚焦在传感网。《国家中长期科学与技术发展规划（2006—2020年）》和“新一代宽带移动无线通信网”重大专项中均将“传感网”列入重点研究领域。经过长期艰苦努力，我国相关机构和企业攻克了大量关键技术，取得了国际标准制定的重要话语权，具备了一定的发展传感网的产业基础，在电力、交通、安防等相关领域的应用也初见成效。

目前我国传感网标准体系已形成初步框架，向国际标准化组织提交的多项标准提案被采纳，传感网标准化工作已经取得积极进展。总的来说，我国在物联网研发上的主要动向如下。

- 江苏省政府与中国移动共同推进TD和传感网基地建设。
- 2009年9月，中国传感网标准工作组成立。
- 2009年9月，举办“感知中国”高峰论坛，探讨如何打造中国传感网产业。
- 2009年9月，无锡市与北京邮电大学就传感网技术签署合作协议，合作建设传感网技术研究院。
- 2009年11月，国家批准无锡建立“国家传感网创新示范区”，使无锡成为中国物联网研究中心。同年，中关村物联网产业联盟成立，成员包括中国移动、清华同方股份有限公司、北京邮电大学、中科院软件所、北京交通委信息中心等十二家单位，囊括了政府、院校和企业。
- 2015年5月，经李克强总理签批，国务院印发了《中国制造2025》，部署全面推进实施制造强国战略，目标是实现长期制约制造业发展的关键共性技术突破，提升我国制造业的整体竞争力。这是我国实施制造强国战略第一个十年的行动纲领。

物联网的市场前景广阔，效益巨大，据预测，到2020年，物物互联业务将达到人人通信业务的30倍，物联网已成为当前各国科技和产业竞争的焦点。

1.1.5 物联网关键技术

1. 感知层关键技术

物联网对事物的感知是以各种信息采集技术为基础的，这些技术主要有RFID、无线传感器网络、二维码、ZigBee、蓝牙、GPS定位等。其中感知层最具代表性的就是RFID和无线传感器网络技术。

1）RFID技术

RFID是英文Radio Frequency Identification的缩写，即射频识别，也称电子标签。它通过无线射频识别不同的目标，利用无线传输技术来存储和检索数据。RFID是一种非接触式的自动识别技术，不需要识别系统与目标有物理、机械或光学的接触，识别工作也无须人为干预，可以工作在较为恶劣的环境中。目前，RFID已经广泛应用于工业自动化、办公自动化、物流、交通管理等多个领域。

RFID 的主要核心部件是阅读器和电子标签。RFID 技术通过相距几厘米到几米甚至几十米距离的阅读器发射的无线电波来读取电子标签内部存储的信息，识别标签所代表的物品信息、状态信息等。一个典型的射频识别系统一般由电子标签、阅读器和计算机系统三部分组成，如图 1-2 所示。



图 1-2　RFID 系统组成

① 电子标签：由耦合元件及芯片组成，每个标签具有唯一的电子编码，附着在需要标识的物体上以识别目标对象，主要由具有模拟、数字记忆功能的芯片，以及依不同频率、应用环境而设计的天线所组成。标签分为无源标签（被动标签）和有源标签（主动标签）两种，无源标签由阅读器的电磁波提供能量，而有源标签自身有电池供电。

② 阅读器：RFID 阅读器通过天线与电子标签进行无线通信，它主要由射频收发单元、模/数转换模块、中央处理单元及天线组成。天线在标签和阅读器间收发信号，中央处理单元用于完成信息的转换和处理。阅读器主要实现标签信息的读取和写入功能。

③ 计算机系统：计算机系统用作后台控制系统，通过有线或无线的通信方式与阅读器相连接，利用阅读器获取标签内部的信息，对读取的数据进行筛选、处理和后台控制，同时也可以根据阅读器的需要查询自身存储的信息，完成标签的认证、信息写入功能。

2）无线传感器网络技术

无线传感器网络（Wireless Sensor Networks，WSNs）是随着微电子机械系统（Micro-Electro-Mechanism System，MEMS）、计算机、通信、自动控制和人工智能等学科的飞速发展而产生的一种新型的测控网络。它是由部署在目标监测区域内大量的价格低廉的微型传感器节点组成，利用无线通信方式形成的一个多跳的自组织的网络系统，其主要目的是感知、采集和处理网络覆盖区域中被感知目标的信息，并发送给观察者或用户。感知目标、传感器节点和观察者是无线传感器网络的三要素。无线传感器网络以一种“无处不在”的计算理念，成为连接物理世界、信息世界和人类社会的桥梁，被广泛应用于环境监控、工业控制、智能家居、国防和公共安全等多个领域。

无线传感器网络通常包括传感器节点（Sensor Node）和汇聚节点（Sink Node，也称基站）。节点可以通过人工布置、飞机抛撒或弹射等方式大量地部署在监测区域内，通过自组织的方式构成无线通信网络，彼此相互协作，感知、采集被监测目标的信息，并通过多跳转发的方法将感知信息经由汇聚节点或基站发送给用户或远程管理中心。同时，用户或远程管理中心也可以对网络内部的节点进行监测和控制。图 1-3 给出了一个典型的无线传感器网络体系结构，其中包括分布式的传感器节点、汇聚节点和管理中心。其中传感器节点在网络中主要负责信息的采集、转发；汇聚节点收集传感器节点采集的信息，并将信息进行协议转换后通过卫星、互联网等其他通信系统发送给管理中心；用户通过管理中心获取所需要的信

息，并经由管理中心和汇聚节点实现对网络的控制。在传感器网络规模较大的情况下，可以通过分层、分簇的方式实现对网络管理的简化。

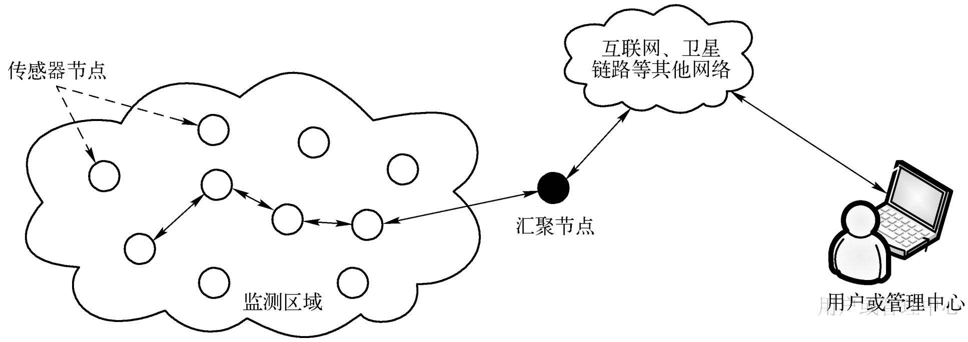


图 1-3　无线传感器网络体系结构

2. 网络层关键技术

网络层是物联网的神经中枢，完成物联网信息的传输和数据的交换。物联网的网络层主要由移动通信网络和互联网组成，从而实现物理世界、虚拟世界、数字世界和人类社会的信息交互。物联网典型的关键技术主要有移动通信技术和计算机网络技术。这里将移动通信技术和计算机网络技术看成两个技术群的集合，它们分别包含很多具体技术手段，具体不再赘述。

1）移动通信技术

移动通信系统包括蜂窝通信、集群调度、无绳电话、无线寻呼等，与固定通信相比，移动通信能够克服通信终端位置对用户的限制，快速、方便地实现信息传输。移动通信系统的一般体系结构如图 1-4 所示。

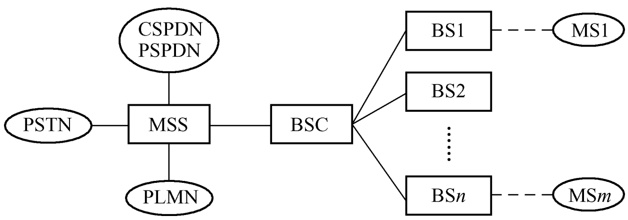


图 1-4　移动通信系统的基本组成

移动通信系统一般由交换系统（MSS）、基站控制器（BSC）、基站（BS）和移动台（MS）组成。移动台是移动终端，如手机等。基站是与移动台通信的固定设备，负责其区域内移动台的通信；基站与移动交换中心通过无线或有线的方式交换信息；移动交换中心与公共电话网相连。公共陆地移动通信网（PLMN）由多个移动通信系统通过数字传输线路互连而成。每个移动通信系统均连接当地的固定电话网络（PSTN）、电路交换公共数据网（CSPDN）、分组交换公共数据网（PSPDN）等。

随着移动通信技术的发展，目前第三代移动通信系统已经普遍商用，主要技术标准有 WCDMA、CDMA2000、TD－SCDMA 三种，能够为用户提供快速便捷的网络服务。同时，

LTE作为3GPP长期演进项目，被看成4G技术，目前很多国家都已经商用，我国也已经广泛应用。

2）计算机网络技术

计算机网络，是指将地理位置不同的具有独立功能的多台计算机及其外部设备，通过通信线路连接起来，在网络操作系统、网络管理软件及网络通信协议的管理和协调下，实现资源共享和信息传递的计算机系统。计算机网络的发展经历了面向终端的单级计算机网络、计算机网络对计算机网络和开放式标准化计算机网络三个阶段。

从地理范围上可以把计算机网络划分为局域网、城域网、广域网和互联网四种。局域网一般来说只能是一个较小区域，城域网是不同地区的网络互连。当然，这种划分并没有严格意义上地理范围的区分，只是一个定性的概念。

① 局域网（LAN），这是我们最常见、应用最广的一种网络，主要采用IEEE 802.3和IEEE 802.11标准。现在，局域网随着整个计算机网络技术的发展和提高得到了充分的应用和普及，几乎每个单位都有自己的局域网，有的甚至家庭中都有自己的小型局域网。很明显，所谓局域网，就是在局部地区范围内的网络，它所覆盖的地区范围较小，网络所涉及的地理距离上一般来说可以是几米至10千米以内。局域网一般位于一个建筑物或一个单位内，不存在寻径问题，不包括网络层的应用。

② 城域网（MAN）一般来说是在一个城市，但不在同一地理小区范围内的计算机互连。这种网络的连接距离可以为10～100千米，它采用的是IEEE 802.6标准。MAN与LAN相比扩展距离更长，连接的计算机数量更多，在地理范围上可以说是LAN网络的延伸。在一个大型城市或都市地区，一个MAN网络通常连接着多个LAN网。例如连接政府机构的LAN、医院的LAN、电信的LAN、公司企业的LAN等。由于光纤连接的引入，使得MAN中高速的LAN互连成为可能。

③ 广域网也称远程网，所覆盖的范围比城域网更广，它一般是在不同城市之间的LAN或MAN网络互连，地理范围可从几百千米到几千千米。因为距离较远，信息衰减比较严重，所以这种网络一般要租用专线，通过IMP（接口信息处理）协议和线路连接起来，构成网状结构，解决循径问题。这种城域网因为所连接的用户多，总出口带宽有限，所以用户的终端连接速率一般较低，通常为9.6kbps～45Mbps，如邮电部的CHINANET、CHINAPAC和CHINADDN。

④ 互联网又称“因特网”。在互联网应用如此发展的今天，它已是我们每天都要打交道的一种网络，无论从地理范围，还是从网络规模来讲它都是最大的一种网络。从地理范围来说，它可以是全球计算机的互连，这种网络的最大的特点就是不定性，整个网络的计算机每时每刻随着人们网络的接入在不断变化。当你连在互联网上时，你的计算机可以算是互联网的一部分，但一旦当你断开互联网的连接，你的计算机就不属于互联网了。但它的优点也是非常明显的，就是信息量大、传播广，无论身处何地，只要连上互联网你就可以对任何连网的用户发出你的信函和广告。因为这种网络的复杂性，所以这种网络实现的技术也是非常复杂的，目前计算机网络技术的研究主要集中在互联网领域。

3. 应用层关键技术

应用层包含一个应用支持子层，感知数据的管理与处理技术是其中的一个核心技术，

包括传感网数据的存储、查询、分析、挖掘、理解及基于感知数据决策和行为的理论和技术。

1）云计算

云计算是一种商业计算模型。它将计算任务分布在由大量计算机构成的资源池上，使各种应用系统能够根据需要获取计算能力、存储空间和各种软件服务。“云”的思想核心就是资源的集中共享，即将所有资源集中在一个先进的平台上，使资源在统一平台的统一管理下，具备最大的灵活性和利用效率，最终以服务的形式提供给用户。云计算平台作为海量感知数据的存储、分析平台，也是物联网支撑技术的重要组成部分，是应用层众多应用的基础。

如今，每个人都能够从互联网中找到自己所需要的东西。像互联网一样，物联网的世界也一定会应用无时不有，价值无处不在。云计算作为一种新的网络基础设施交付和使用模式，将在一定程度上改变传统思维，促进物联网和互联网更好发展。

2）大数据分析

移动互联网、物联网的出现带来了数据的爆炸性增长，这些数据不但体量大、增长速度快，还存在种类多样、结构多样、价值密度低等特点，如此庞大、复杂的数据利用传统的数据存储、分析方法处理已经力不从心，大数据技术应运而生，其理论核心就是数据挖掘算法，各种数据挖掘的算法基于不同的数据类型和格式才能更加科学地呈现出数据本身具备的特点，挖掘出公认的价值。

从技术上看，大数据与云计算的关系就像一枚硬币的正反面一样密不可分。大数据必然无法用单台的计算机进行处理，必须采用分布式架构。它的特色在于对海量数据进行分布式数据挖掘。但它必须依托云计算的分布式处理、分布式数据库、云存储、虚拟化技术。物联网应用层利用经过大数据分析处理的信息，能够为用户提供丰富的特定服务。

1.2 物联网安全

物联网是一种虚拟网络与现实世界实时交互的新型系统，其无处不在的数据感知、以无线为主的信息传输、智能化的信息处理能够有效提高社会效率，但也日益引起大众对信息安全和隐私保护问题的关注，特别是暴露在公开场所之中的无线信号很容易被窃取、干扰，这将直接影响到物联网体系的安全。因此，物联网安全技术将成为物联网应用的重要支撑，对于安全技术的研究也有着重要的意义。

1.2.1 物联网安全特点及面临的安全挑战

物联网系统的安全和一般 IT 系统的安全基本一样，主要有八个尺度：读取控制、隐私保护、用户认证、不可抵赖性、数据保密性、通信层安全、数据完整性、随时可用性。前四项主要处在物联网三层架构的应用层，后四项主要位于传输层和感知层。其中，隐私权和可信度（数据完整性和保密性）问题在物联网体系中尤其受到关注。

根据物联网自身的特点，物联网除了面对移动通信网络的传统网络安全问题之外，还存在一些特有的安全问题，这主要是由于物联网大量的异构终端、分散的网络部署、灵活的接

入方式、多样的用户需求等特点决定的。这些特殊的安全问题主要有以下几个方面。

1. 物联网设备的本地安全问题

由于物联网设备常常部署在环境恶劣、危险、复杂的地域，如火山、海底、战场等，因此通常设备单独工作，很难进行及时维护。在这种无人场景下，物联网设备很容易暴露在攻击者面前，攻击者也可以轻易地获取设备，对其进行破坏、破解、仿制等攻击。例如 RFID 标签已经大量使用，这些标签结构简单、成本低廉，很容易被所有者忽略并被攻击者获取，从而进行破解、仿制等攻击。

2. 感知层网络的传输与信息安全问题

物联网感知层设备极其丰富、结构多种多样、标准也层出不穷，且感知层设备由于数量庞大，要求成本低廉，因此其自身结构也相应比较简单。这就使得感知层设备首先自身能力有限，无法执行复杂的安全操作，如 RFID 标签仅能存储有限的数据和执行简单的安全算法；同时，由于感知层需要传输的数据多种多样，采用的标准又难以统一，因此无法为感知层提供统一的安全保护体系。另外，由于感知层大多采用无线传输技术，通信数据很容易被攻击者窃听并进行分析，因此物联网感知层信息安全问题较为突出。

3. 核心网络的传输与信息安全问题

物联网以互联网、移动通信网络为核心，这些网络具有相对完整的安全保护能力，但由于物联网中节点数量庞大，且以集中方式存在，因此会导致在数据传输时，常常出现大量的数据突发传输，造成网络拥堵，产生网络的拒绝服务攻击。另外，现有通信网络的安全架构是以人的通信特点进行设计的，是否适用于机器之间的通信还有待进一步验证，现有的安全机制可能会割裂物联网机器间的逻辑关系。

4. 物联网应用的安全问题

由于物联网设备可能事先部署，而后连接网络，而且物联网设备多处于无人管理、动态变化的状态，因此如何对物联网设备进行远程管理和配置就成为一个问题。另外，庞大的异构节点、网络的管理也需要一个强大而统一的管理平台，否则独立平台会因为物联网设备的异构特点而同样复杂多样，被应用所掩盖，这就使得网络的安全管理成为难题，且异构管理平台间的安全通信又成为新的安全问题。

针对上述安全问题，物联网发展的中、高级阶段面临如下五大特有（在一般 IT 安全问题之上）的信息安全挑战。

① 有线长、短距离和无线长、短距离四类网络相互连接组成的异构、多级、分布式的融合网络导致统一的网络安全体系难以实现桥接和过渡。

② 设备大小不一、存储能力不同、处理能力不同导致信息安全操作、传递、处理难以统一。

③ 物联网设备可能无人值守、丢失，也可能处于运动状态，连接具有间歇性，可信度差，这种种因素导致了物联网信息安全系统的设计和实施复杂度增加，很难实现。

④ 在保证一个物联网智能终端能够被大量的其他终端、识别设备认知和接受的同时，还要保证数据传输的安全可靠及保证设备所具有的隐私性，难度较高。

⑤ 物联网设备广泛应用，感知的信息丰富多样，即便在信息加密的条件下也能够通过数据流向、流量、时间、地址等元数据对通信进行挖掘，从而泄露隐私，隐私保护面临前所未有的挑战。

物联网的安全特点及其安全面临的挑战为物联网安全技术的研究提供了广阔的空间，也有着较高的难度，目前已成为国内外物联网研究的焦点。

1.2.2　物联网面临的安全威胁与安全目标

1. 物联网面临的安全威胁

人们可以通过物联网感知各方面的信息，同时也可以通过物联网实现各种应用。现实世界中的物体都连接成网络，我们可以远程感知和控制类似家电、交通、能源及金融等设施和服务。在物联网提供这些便利的同时，人们也对物联网的强大表现出了忧虑。物联网在网络的每个层次上都存在威胁：感知层方面有终端设备的物理安全、信息的传输安全、隐私泄露等问题，网络层方面有数据破坏、身份假冒及信息泄露等安全问题，应用层方面有身份假冒、非法接入、越权操作等安全问题。下面将从物联网各个层次上分别对物联网面临的安全威胁进行简单的介绍，由于网络层和应用层的攻击在当前互联网中已经有所体现，所以这里将以感知层面临的安全威胁作为重点阐述。

1）物联网感知层安全威胁

物联网感知层的任务是全面感知外界信息，该层的典型设备包括 RFID 设备和传感器网络设备，因此感知层安全威胁主要从这两个方面介绍。

（1）RFID 系统的安全缺陷与攻击行为

RFID 所面临的安全问题比传统网络要严峻得多，这主要是 RFID 系统自身属性所带来的安全缺陷造成的，这种安全属性不仅仅表现在 RFID 产品的处理能力和安全措施的有限性，更主要是 RFID 技术本身就包含了比计算机和网络更多、更容易泄密的不安全因素。这些属于 RFID 技术本身的安全缺陷如下。

① 标签本身的访问缺陷。RFID 标签一般成本受限，很难具有保证自身安全的能力，因此非法用户完全可以使用合法的阅读器或自制的阅读器与标签通信，获取或修改标签内部存储的数据。

② 通信链路的安全问题。RFID 一般采用近距离无线通信技术，无线信号面向空间开放，攻击者很容易进行非法侦听，如攻击者可以非法截取通信数据、可以通过发射干扰信号阻塞通信链路、可以大量发送读/写请求，瘫痪设备，还可以利用假冒标签欺骗合法阅读器使其提供虚假信息。

③ 阅读器内部的安全风险。在阅读器中，除了中间件被用来完成数据的传输选择、时间过滤和管理之外，只能提供用户业务接口，而不能提供让用户自行提升安全性能的接口。

④ 标签的安全风险。RFID 标签电路结构简单，存储单元小，处理能力弱，攻击者很容易实现对标签的捕获和分析，并复制大量非法的标签。

⑤ 标签的隐私泄露。RFID 承载着众多攻击者感兴趣的信息，除了对标签实施攻击外，攻击者还可以通过对标签的分析来实现目标隐私的获取。

由此可见，RFID 所遇到的安全问题相比普通计算机网络而言并不简单，甚至更为复杂，特别是在电子标签上，计算能力和可编程能力都被标签本身的成本所限定，更准确地讲，在一个特定的应用中，标签的成本越低，它的计算能力也就越弱，实施安全操作的难度也就越高。目前，RFID 系统常见的攻击行为主要有 4 种类型，见表 1-1。

表 1-1　RFID 的攻击分类

攻 击 分 类	具体实施过程
电子标签数据的捕获攻击	未授权攻击者进入授权的读写器时，仍然能够设置阅读器与某一特定的电子标签通信，并读取甚至修改标签上的信息，电子标签的数据就会受到攻击
电子标签与阅读器间的通信入侵	通过非法阅读器截获合法标签与阅读器间的数据、第三方堵塞数据传输、伪造标签发送数据等方法来干扰电子标签和阅读器间的正常通信
入侵阅读器内部数据	攻击者设法获得阅读器存储在阅读器内存中的数据来实施攻击
主机系统入侵	通过后台主机系统的入侵来获得更多关于标签及其他方面的信息

（2）传感器网络的安全威胁

传感器网络是物联网感知层的重要组成部分，无论是在研究还是应用领域都极为重视。由于传感器网络节点众多，部署范围广，维护困难，因此对网络的安全保护比较困难，相反，攻击者则很容易实施攻击。图 1-5 给出了传感器网络的自身安全特点，由于这些特点使得传感器网络很容易受到攻击。表 1-2 给出了传感器网络各层次的攻击及防御方法。

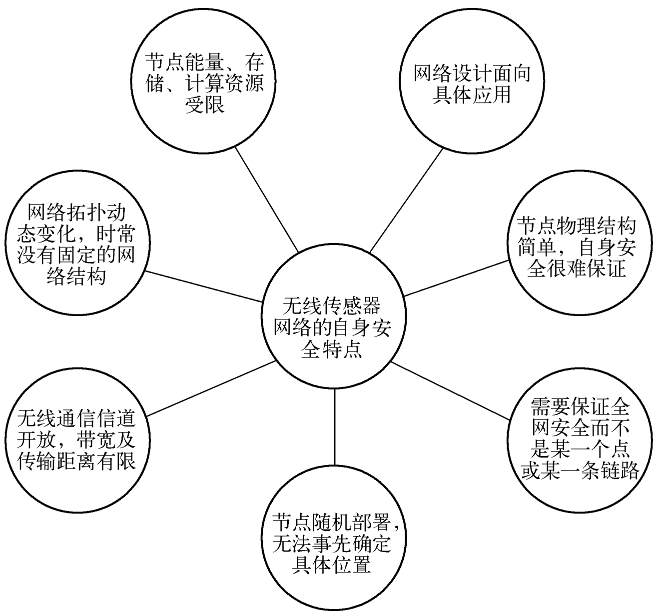


图 1-5　无线传感器网络的自身安全特点

表 1–2　传感器网络攻击方法层次划分及其防御方法

传感器网络层次	攻 击 方 法	防 御 方 法
物理层	监听攻击	访问控制、加密、硬件防篡改技术
	拥塞攻击	
	设备篡改攻击	
MAC 层	传输控制	恶意行为检测、身份保护等
	身份篡改	
	非公平竞争	
网络层	虚假路由	路由访问控制、虚假路由信息检测、虫洞检测
	数据包复制	
	黑洞攻击	
	污水池攻击	
	选择性转发攻击	
	蠕虫攻击	
应用层	时钟偏差	数据完整性保护、数据机密性保护等
	选择性消息转发	
	数据聚合失真	

从表 1–2 中可以看出，传感器网络的物理层由于涉及节点的物理实体、通信信号的检测、编码、调制、频率等方面，因此在物理层上攻击者可以实施设备篡改攻击、监听攻击和拥塞攻击。当攻击者获得网络使用的频段、通信速率、编码方式、调制方式等信息后，一方面可以使用无线接收设备侦听通信信道，从而获得通信数据，另一方也可以不间断地在信道上发送无用信号，对信道实施阻塞。同时，由于传感器网络节点难以维护，因此攻击者一旦捕获节点，并对节点进行分析、破解，就有可能获得节点的工作原理和存储的相关秘密信息，对网络实施破坏。

传感器网络的 MAC 层主要用来协调无线传输的公平性和效率，在 MAC 层协议中使用了很多经典的节点信息交换控制包来保证信息能够在一定阶段内获得相关的通信权限。传输控制、身份篡改和非公平性竞争是 MAC 层的常见攻击方式。现有广泛应用的 MAC 协议中，信息的传输都是在一定规则下进行的。攻击者可以通过竞争手段来实现传输的控制，也可以通过改变数据包的优先级来实现非公平性竞争；另外，攻击者还可以通过改变自身的身份标识来实现身份的篡改。

传感器网络的网络层主要实现数据和消息的路由。攻击者可以通过获得的路由信息实施网络层攻击，攻击的方法和手段更加多样。例如攻击者可以通过散布虚假路由信息改变数据传输的路径，使得信息不可达或改变网络负载均衡；可以通过复制网络中已经发送的数据形成泛洪攻击，耗费网络资源，减少网络寿命；可以散布更短的通信路径，形成路由黑洞，吸引网络数据；可以针对数据包进行选择性的转发，影响网络通信。

在应用层，攻击者能够根据传感器网络中的各种应用的特点进行攻击。时钟偏差攻击主要针对那些需要同步操作的传感器网络，通过传播虚假的时间信息来使网络节点无法进行时间同步；消息的选择性转发攻击可以针对消息的具体内容来进行有选择的信息传

输；攻击者在数据传输到基站前，可能通过篡改数据，改变网络数据聚合结构，实施聚合攻击。

针对传感器网络的攻击及防御手段的研究已经非常普遍，取得了大量的研究成果，不过传感器网络由于其网络特点及安全问题难度大，仍然是网络研究的热门领域。

2）物联网网络层安全威胁

物联网网络层主要用于把感知层获得数据安全、可靠地传输到应用层，或者将应用层的指令数据发送到感知层。网络层的基础设施主要包括互联网、移动通信网络和一些其他专用网络，如电力网、广电网等。在信息传输过程中，可能经过一个或多个不同架构的网络。与现有网络环境的安全性相比，由于网络层中不同架构的网络需要互连互通，物联网在跨网络架构的安全方面面临更大的挑战。但总体而言，从目前来看，已有的网络攻击手段仍然是物联网面临的主要威胁，比较典型的安全威胁如下。

（1）DoS 攻击、DDoS 攻击

在物联网发展过程中，目前的互联网或下一代互联网将是物联网网络层的核心载体，多数信息要经过互联网进行传输。互联网遇到的 DoS 和 DDoS 攻击仍然存在，因此需要有更好的防范措施和灾难恢复机制。考虑到物联网所连接的终端设备性能和对网络需求的巨大差异，对网络攻击的防护能力也会有很大差别，因此很难设计通用的安全方案，而应针对不同的网络和需求设计不同的安全防范措施。

（2）假冒攻击、中间人攻击等

在网络层，异构网络的信息交换将成为安全性的脆弱点，特别是在网络认证方面，难免存在假冒攻击、中间人攻击和其他类型的攻击，如异步攻击、合谋攻击等，这些攻击都要有更高级的安全防护措施。

（3）数据传输安全

数据传输安全包括数据传输的机密性、完整性和数据流机密性。在网络层的数据传输过程中，必须保证不泄露数据内容，数据不被非法篡改或非法篡改的数据容易被检测出，某些应用场景还需要对数据流量信息进行保密。

3）物联网应用层安全威胁

应用层设计的是综合的或有个体特性的具体应用业务，它所涉及的某些安全问题通过前面几个逻辑层的安全解决方案可能仍然无法解决。在这些问题中，隐私保护就是典型的一种。无论感知层、传输层还是处理层，都不涉及隐私保护的问题，但它却是一些特殊应用场景的实际需求，即应用层的特殊安全需求。物联网的数据共享有多种情况，涉及不同权限的数据访问。此外，在应用层还将涉及知识产权保护、计算机取证、计算机数据销毁等安全需求和相应技术。

应用层的安全挑战和安全需求主要来自于下述几个方面：

① 如何根据不同访问权限对同一数据库内容进行筛选；

② 如何提供用户隐私信息保护，同时又能正确认证；

③ 如何解决信息泄露追踪问题；

④ 如何进行计算机取证；

⑤ 如何销毁计算机数据；

⑥ 如何保护电子产品和软件的知识产权。

2. 物联网的安全目标

从技术角度来说，物联网信息安全的目标与其他网络类似，主要表现在物联网信息及系统的可靠性、可用性、保密性、完整性、不可抵赖性、可控性、隐私性等方面。

1） 可靠性

可靠性是指网络信息系统能够在规定条件下和规定的时间内完成规定功能的特性。可靠性是系统安全的最基本要求之一，是所有网络信息系统的建设和运行目标。网络信息系统的可靠性测度主要有三种：抗毁性、生存性和有效性。抗毁性是指系统在人为破坏下的可靠性，生存性是在随机破坏下系统的可靠性，有效性是一种基于业务性能的可靠性。可靠性主要表现在硬件可靠性、软件可靠性、人员可靠性、环境可靠性等方面。

2） 可用性

可用性是网络信息可被授权实体访问并按需求使用的特性。即网络信息服务在需要时，允许授权用户或实体使用的特性，或者是网络部分受损或需要降级使用时，仍能为授权用户提供有效服务的特性。可用性是网络信息系统面向用户的安全性能。网络信息系统最基本的功能是向用户提供服务，而用户的需求是随机的、多方面的，有时还有时间要求。可用性一般用系统正常使用时间和整个工作时间之比来度量。

3） 保密性

保密性是网络信息不被泄露给非授权的用户、实体或过程，或供其利用的特性。即防止信息泄露给非授权个人或实体，信息只为授权用户使用的特性。保密性是在可靠性和可用性基础之上，保障网络信息安全的重要手段。常用的保密技术包括：防侦收、防辐射、信息加密、物理保密。

4） 完整性

完整性是网络信息未经授权不能进行改变的特性。即网络信息在存储或传输过程中保持不被偶然或蓄意地删除、修改、伪造、乱序、重放、插入等而破坏和丢失的特性。完整性是一种面向信息的安全性，它要求保持信息的原样，即信息的正确生成、存储和传输。完整性与保密性不同，保密性要求信息不被泄露给未授权的人，而完整性则要求信息不致受到各种原因的破坏。

5） 不可抵赖性

不可抵赖性也称不可否认性，在网络信息系统的信息交互过程中，确信参与者的真实同一性。即所有参与者都不可能否认或抵赖曾经完成的操作和承诺。利用信息源证据可以防止发信方不真实地否认已发送信息，利用递交接收证据可以防止收信方事后否认已经接收的信息。

6） 可控性

可控性是对网络信息的传播及内容具有控制能力的特性，能够及时对不安全、非法操作或信息进行屏蔽，保证网络健康发展。

7）隐私性

隐私性是网络对个人、机构等实体不愿意被外部世界知晓的信息保护能力。隐私性要求较高，不但要保证消息的保密性，还要保证用户的各类操作及通信能够不被攻击者所分析并推理出相应的信息。

1.2.3　物联网安全技术框架

结合物联网面临的安全威胁与安全目标，为满足物联网各层次的安全需求，保障网络安全，本书提出了物联网安全技术框架，如图 1-6 所示，它包括技术因素和非技术因素两部分，其中技术因素部分主要包括感知层安全技术、网络层安全技术和应用层安全技术。

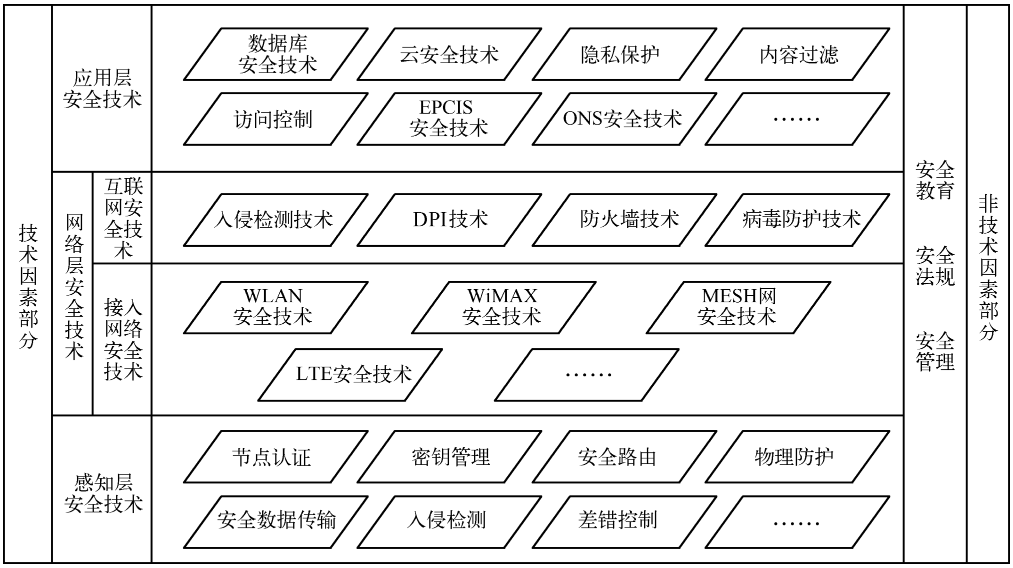


图 1-6　物联网安全技术框架

1. 感知层安全技术

感知层安全技术的主要目的是保护物联网感知层节点、网络的物理安全，保障节点、网络可靠运行，保证数据通信的机密和完整。相应的安全技术主要包括感知层节点物理防护技术、安全认证技术、密钥管理技术、网络安全路由技术、差错控制技术、信息编码技术、入侵检测技术等内容。

① 物理防护技术主要是防止感知层节点在被攻击者捕获后从中获取敏感信息。由于物联网感知层节点设备通常比较简单，很容易被攻击者破解，因此物理防护技术主要是通过必要加密手段、自毁命令等方式实现敏感数据的保密。

② 安全认证技术主要是实现用户对目标节点、节点之间、节点对通信数据的安全认证，确认通信目标的合法性和通信数据的来源可靠性。物联网是物与物之间的网络，因此实现节点间的安全认证，防止非法节点接入网络对于保证网络安全有着非常重要的作用。

③ 密钥管理技术主要是保障网络通信的机密性，防止攻击者对窃听到的数据实现破解。密钥管理是保障网络安全最重要的安全技术手段，在各类网络中被普遍使用。

④ 针对传感器网络等感知层网络，安全路由技术主要用于保障网络路由的建立、维护和运行的安全，确保信息传输路径的高效、可靠，防止攻击者通过路由信息破坏网络的正常运行。

⑤ 差错控制技术主要防止数据在通信过程中因为环境等原因造成的数据传输错误，实现数据的纠错和检错，提高通信效率。

⑥ 信息编码技术主要是利用信道编码、网络编码等手段对通信数据进行信息处理，保证通信的高效性和安全性。

⑦ 入侵检测技术主要是利用协同感知、流量分析、行为分析等方法实现攻击的检测和分析。

2. 网络层安全技术

网络层安全技术的主要目的是保护物联网网络通信的安全。实现感知层数据对网络的安全接入，保障网络运行的安全、可靠，确保不同网络互连互通的安全。相应的安全技术主要包括接入网络安全技术和互联网安全技术。

接入网络安全技术主要有 WLAN 安全技术、WiMAX 安全技术、MESH 网安全技术及 4G 安全技术等。这些安全技术能够有效确保感知层数据对网络接入的安全，确保网络层边界及信息来源的可靠。

互联网是物联网的核心，因此物联网网络层安全很大程度上仍然是依靠传统的互联网、计算机网络安全技术，这些技术主要有防火墙技术、入侵检测技术、DPI 技术、恶意代码检测技术、病毒防护技术等。

3. 应用层安全技术

应用层安全技术主要是确保物联网各类应用及业务的安全，主要包括保证数据安全的数据库安全技术、保证服务安全的云计算安全技术、保证信息内容安全的内容过滤技术、保证业务安全的访问控制技术、保证物联网信息服务系统安全的 EPCIS 安全技术等。物联网应用层涉及的具体应用及业务类型多种多样，因此对应的安全技术也种类繁多。

4. 非技术因素

物联网的安全问题不仅仅是技术问题，同样也涉及许多非技术性的因素，以下因素就是很难通过技术手段实现的。

① 安全教育。通过安全教育让用户意识到信息安全的重要性，从而正确地使用物联网所提供的各种服务，减少敏感信息的泄露。

② 安全管理。找到信息系统安全方面的薄弱环节并对其进行强化管理，可以提高系统的整体安全程度。这些薄弱环节包括资源管理、物理安全管理和人力安全管理等多个方面。

③ 安全法规。制定信息安全法律、法规，从法律和制度层面上规定什么信息可以自由流动、什么信息不能流动。明确物联网内容传播过程中各个环节、相关机构的责任和义务能够有效地防止信息安全事件的发生。

参考文献

［1］王汝传，孙力娟．物联网技术导论［M］. 北京：清华大学出版社，2011.
［2］王喜富．物联网与物流信息化［M］. 北京：电子工业出版社，2011.
［3］张铎．物联网大趋势［M］. 北京：清华大学出版社，2010.
［4］王良民，熊书明．物联网工程概论［M］. 北京：清华大学出版社，2011.
［5］孙利民，李建中，等．无线传感器网络［M］. 北京：清华大学出版社，2005.
［6］武奇生，刘盼芝．物联网技术与应用［M］. 北京：机械工业出版社，2011.
［7］王毅，镇维，等．物联网技术及应用［M］. 北京：国防工业出版社，2011.
［8］沈玉龙，裴庆祺，等．无线传感器网络安全技术概论［M］. 北京：人民邮电出版社，2010.
［9］王汝林，王小宁，等．物联网基础及应用［M］. 北京：清华大学出版社，2011.
［10］郭渊博，杨奎武，等．ZigBee 技术与应用［M］. 北京：国防工业出版社，2010.
［11］郭渊博，杨奎武，等．无线局域网安全、设计及实现［M］. 北京：国防工业出版社，2010.
［12］徐小涛，杨志红，等．物联网信息安全［M］. 北京：人民邮电出版社，2012.

第 2 章　密码技术基础

密码学是现代信息安全技术的基础，加密、数字签名、认证等都与密码技术有着密切的关系。通过密码算法，用户不仅可以保护自己的敏感数据，还可以进行安全可靠的网络交易、网络支付，建立网络上的信任关系。本章主要介绍简单的密码学理论，以方便读者对后续内容的理解。

2.1　密码学概述

密码学是一门古老的科学，自古以来密码主要用于军事、政治、外交等重要部门，因而密码学的研究工作也是秘密进行的。密码学的知识和经验主要掌握在军事、政治、外交等保密机关，不便公开发表。然而随着计算机科学技术、通信技术、微电子技术的发展，计算机和通信网络的应用进入了人们的日常生活和工作中，出现了电子政务、电子商务、电子金融等必须确保信息安全的系统，使得民间和商界对信息安全保密的需求大增。总而言之，在密码学形成和发展的历程中，科学技术的发展和战争的刺激起着积极的推动作用。

密码技术形成一门新的学科是在 20 世纪 70 年代，这是受计算机科学蓬勃发展和推动的结果。密码学的理论基础之一是 1949 年 Shannon 发表的《保密系统的通信理论》（The Communication Theroy of Secrecy System），这篇文章发表了 30 年后才显示出它的价值。1976 年，W. Diffie 和 M. Hellman 发表了《密码学的新方向》（New Direction Cryptography）一文，提出了适应网络保密通信的公钥密码思想，开辟了公开密钥密码学的新领域，掀起了公钥密码研究的序幕。各种公钥密码体制被提出，特别是 1978 年 RSA 公钥密码体制的出现，在密码学史上是一个里程碑。同年，美国国家标准局正式公布实施了美国的数据加密标准（Data Encryption Standard，DES），宣布了近代密码学的开始。2001 年，美国联邦政府颁布高级加密标准（Advanced Encryption Standard，AES）。随着其他技术的发展，一些具有潜在密码应用价值的技术也逐渐得到了密码学家极大的重视并加以利用，出现了一些新的密码技术，如混沌密码、量子密码等，这些新的密码技术正在逐步走向实用化。

2.1.1　密码体制

研究各种加密方案的科学称为密码编码学（Cryptography），而研究密码破译的科学称为密码分析学（Cryptanalysis）。密码学作为数学的一个分支，是密码编码学和密码分析学的统称，其基本思想是对信息进行一系列的处理，使未授权者不能获得其中的真实含义。

一个密码系统也称密码体制（Cryptosystem），有五个基本组成部分，如图 2-1 所示。

明文：是加密输入的原始信息，通常用 m 表示。全体明文的集合称为明文空间，通常

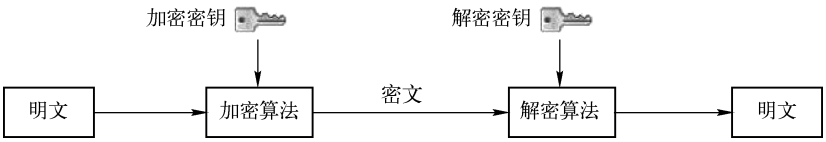


图 2-1　密码系统模型

用 M 表示。

密文：是明文经过加密变换后的结果，通常用 c 表示。全体密文的集合称为密文空间，通常用 C 表示。

密钥：是参与信息变换的参数，通常用 k 表示。全体密钥的集合称为密钥空间，通常用 K 表示。

加密算法：是将明文变成密文的变换函数，即发送者加密消息时所采用的一组规则，通常用 E 表示。

解密算法：是将密文变成明文的变换函数，即接收者解密消息时所采用的一组规则，通常用 D 表示。

加密：是将明文 M 用加密算法 E 在加密密钥 K_e 的控制下变换成密文 C 的过程，表示为 $C=E_{K_e}(M)$。

解密：是将密文 C 用解密算法 D 在解密密钥 K_d 的控制下变换成明文 M 的过程，表示为 $M=D_{K_d}(C)$，并要求 $M=D_{K_d}(E_{K_e}(M))$，即用加密算法得到的密文用一定的解密算法总能够恢复成为原始的明文。

对称密码体制：当加密密钥 K_e 与解密密钥 K_d 是同一把密钥，或者能够相互较容易地推导出来时，该密码体制被称为对称密码体制。

非对称密码体制：当加密密钥 K_e 与解密密钥 K_d 不是同一把密钥，且解密密钥不能通过加密密钥计算出来（至少在假定合理的长时间内）时，该密码体制被称为非对称密码体制。

在密码学中通常假定加密密钥和解密算法是公开的，密码体制的安全性只系于密钥的安全性，这就要求加密算法本身要非常安全。如果提供了无穷的计算资源，依然无法攻破，则称这种密码体制是无条件安全的。除了一次一密之外，无条件安全是不存在的，因此密码系统用户所要做的就是尽量满足以下条件：

- 破译密码的成本超过密文信息的价值；
- 破译密码的时间超过密文信息有用的生命周期。

如果满足上面两个条件之一，则密码系统可以认为是实际上安全的。

2.1.2　密码分类

加密技术除了隐写术以外可以分为古典密码和现代密码两大类。古典密码一般是以单个字母为作用对象的加密法，具有久远的历史；而现代密码则以明文的二元表示作为作用对象，具备更多的实际应用。现将常用密码算法按照古典密码与现代密码归纳，如图 2-2 所示。本书不对古典密码进行介绍，感兴趣的读者可以参考其他密码学教程。

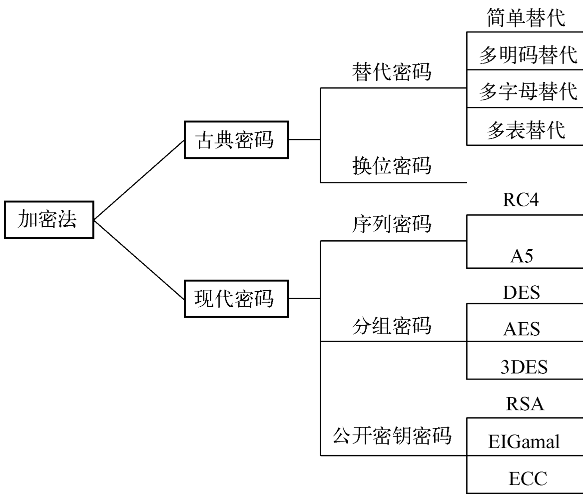


图 2-2　加密法分类图

2.2　分组密码

现代密码学中所出现的密码体制可分为两大类：对称加密体制和非对称加密体制。对称加密体制中相应采用的就是对称算法。在大多数对称算法中，加密密钥和解密密钥是相同的。从基本工作原理来看，古典加密算法最基本的替代和换位工作原理仍是现代对称加密算法最重要的核心技术。对称算法可以分为两类：序列密码（Stream Cipher）和分组密码（Block Cipher），其中绝大多数基于网络的密码应用，使用的是分组密码。

与序列密码每次加密处理数据流的一位或一字节不同，分组密码处理的单位是一组明文，即将明文消息编码后的数字序列划分成长为 L 位的 m 组，每个长为 L 的分组分别在密钥 k（密钥长为 t）的控制下变换成与明文分组等长的一组密文数据文字序列 c。

分组密码算法实际上就是在密钥的控制下，通过某个置换来实现对明文分组的加密变换。为了保证密码算法的安全强度，对密码算法的要求如下：

- 分组长度足够长；
- 密钥量足够多；
- 密码变换足够复杂。

2.2.1　DES

美国国家标准局（NBS）于 1973 年向社会公开征集一种用于政府机构和商业部门的加密算法，经过评测和一段时间的试用，美国政府于 1977 年颁布了数据加密标准（Data Encryption Standard，DES）。DES 是分组密码的典型代表，也是第一个被公布出来的标准算法，曾被美国国家标准局确定为联邦信息处理标准（FIPS PUB 46），使用广泛，特别是在金融领域，曾是密码体制事实上的世界标准。

DES 是一种分组密码，明文、密文和密钥的分组长度都是 64 位，并且是面向二进制的密码算法。DES 处理的明文分组长度为 64 位，密文分组长度也是 64 位，使用的密钥长度为

56 位（实际上函数要求一个 64 位的密钥作为输入，但其中用到的只有 56 位，另外 8 位可以用作奇偶校验或完全随意设置）。DES 是对合运算，它的解密过程和加密相似，解密时使用与加密同样的算法，不过子密钥的使用次序则要与加密相反。DES 的整个体制是公开的，系统的安全性完全靠密钥保密。DES 的整体结构如图 2-3 所示。

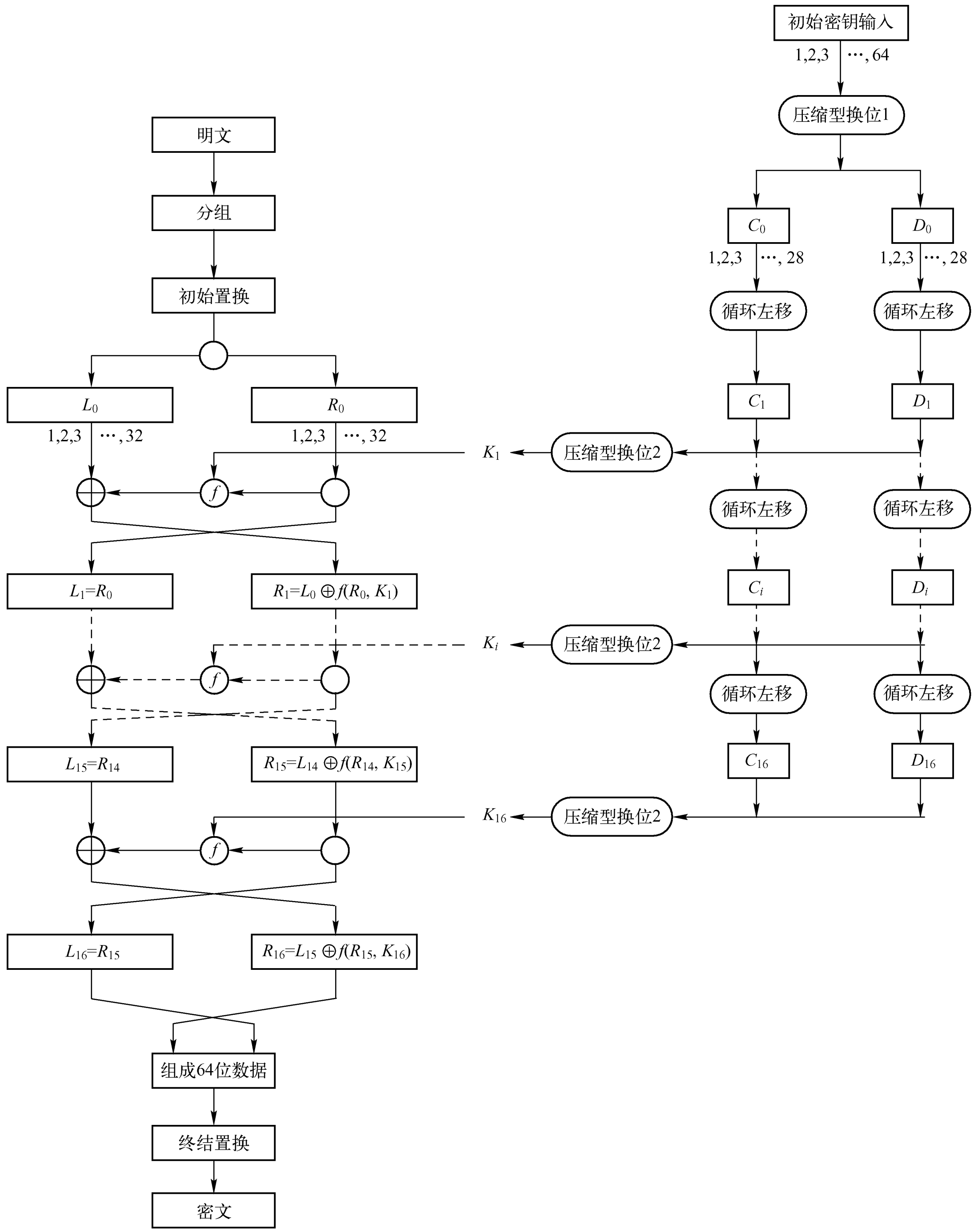


图 2-3　DES 的整体结构

DES 算法的加密过程经过了三个阶段：首先，64 位的明文在一个初始置换 IP 后，比特重排产生了经过置换的输入，明文组被分成右半部分和左半部分，每部分 32 位，以 L_0 和 R_0 表示；第二阶段是对同一个函数进行 16 轮迭代，称为乘积变换或函数 f。这个函数将数据和密钥结合起来，本身既包含换位又包含替代函数，输出为 64 位，其左边和右边两个部分经过交换后得到预输出。

$$\begin{cases} L_i = R_{i-1} \\ R_i = L_{i-1} \oplus f(R_{i-1}, K_i) \end{cases} \quad i = 1, 2, 3, \cdots, L, \cdots, 16$$

最后阶段，预输出通过一个逆初始置换 IP^{-1} 算法就生成了 64 位密文结果。相对应的 DES 解密过程由于 DES 的运算是对合运算，所以解密和加密可以共用同一个运算，只是子密钥的使用顺序不同。解密过程可以用如下公式表示：

$$\begin{cases} R_{i-1} = L_i \\ L_{i-1} = R_i \oplus f(L_i, K_i) \end{cases} \quad i = 1, 2, 3, \cdots, L, \cdots, 16$$

DES 在总体上应该说是极其成功的，但在安全上也有不足之处。

- 密钥太短：IBM 原来的 Lucifer 算法的密钥长度是 128 位，而 DES 采用的是 56 位，显然太短了。1998 年 7 月 17 日美国 EFF（Electronic Frontier Founation）宣布，他们使用一台价值 25 万美元的改装计算机，只用了 56 小时就穷举出一个 DES 密钥。1999 年 EFF 将该穷举速度提升到了 24 小时。
- 存在互补对称性：将密钥的每一位取反，用原来的密钥加密已知明文得到密文分组，那么用此密钥的补密钥加密此明文的补便可得到密文分组的补。这表明，对 DES 的选择明文攻击仅需要测试一半的密钥，穷举攻击的工作量也就相应减半了。

除了上述两点之外，DES 的半公开性也是人们对 DES 颇有微辞的地方。后来虽然推出了 DES 的改进算法，如三重 DES，即 3DES，将密钥长度增加到 112 位或 168 位，增强了安全性，但效率低。

2.2.2 AES

高级加密标准（Advanced Encryption Standard，AES）作为传统对称加密标准 DES 的替代者，于 2001 年正式发布为美国国家标准（FIST PUBS 197）。

AES 采用的 Rijndael 算法是一个迭代分组密码，其分组长度和密钥长度都是可变的，只是为了满足 AES 的要求才限定处理的分组大小为 128 位，而密钥长度为 128 位、192 位或 256 位，相应的迭代轮数 N 为 10 轮、12 轮、14 轮。Rijndael 汇聚了安全性能、效率、可实现性和灵活性等优点，其最大的优点是可以给出算法的最佳差分特性的概率，并分析算法抵抗差分密码分析及线性密码分析的能力。Rijndael 对内存的要求非常低且操作简单，也使它很适合用于受限的环境中，并可抵御强大和实时的攻击。

在安全性方面，Rijndael 加密、解密算法不存在像 DES 里出现的弱密钥，因此在加密、解密过程中，对密钥的选择就没有任何限制；并且根据目前的分析，Rijndael 算法能够有效抵抗现有已知的攻击。

除了前面介绍的分组密码外，还有其他很多的分组密码，如 RC 系列分组密码（包括 RC2、RC5、RC6 等）、CLIPPER 密码、SKIPJACK 算法、IDEA 密码等。国际上目前公开的分组密码不下 100 种，在此不一一介绍。

2.3　公钥密码体制

公钥密码学与其之前的密码学完全不同。首先，公钥密码算法基于数学函数而不是之前的替代和置换。其次，公钥密码学是非对称的，它使用两个独立的密钥。公钥密码学在消息的保密性、密钥分配和认证领域都有着极其重要的意义。

公开密钥密码的基本思想是将传统密码的密钥 k 一分为二，分为加密密钥 K_e 和解密密钥 K_d，用加密密钥 K_e 控制加密，用解密密钥 K_d 控制解密，而且在计算上确保由加密密钥 K_e 不能算出解密密钥 K_d。这样即使将 K_e 公开也不会暴露 K_d，从而不会损害密码的安全。于是可对 K_d 保密，而对 K_e 进行公开，从而在根本上解决了传统密码在密钥分配上所遇到的问题。为了区分常规加密和公开密钥加密两个体制，一般将常规加密中使用的密钥称为秘密密钥（Secret Key），用 K_s 表示；将公开密钥加密中使用的能够公开的加密密钥 K_e 称为公开密钥（Public Key），用 K_U 表示；将加密中使用的保密的解密密钥 K_d 称为私有密钥（Private Key），用 K_R 表示。

根据公开密钥密码的基本思想，可知一个公开密钥密码应当满足以下三个条件：

- 解密算法 D 与加密算法 E 互逆，即对所有明文 M 都有 $D_{K_R}(E_{K_U}(M))=M$；
- 在计算上不能由 K_U 推出 K_R；
- 算法 E 和 D 都是高效的。

满足了以上三个条件，便可构成一个公开密钥密码，这个密码可以确保数据的秘密性。进而，如果还要求确保数据的真实性，则还应该满足第四个条件，即：

对于所有明文 M 都有 $E_{K_U}(D_{K_R}(M))=M$。

如果同时满足以上四个条件，则公开密钥密码可以同时确保数据的秘密性和真实性。此时，对于所有的明文 M 都有 $D_{K_R}(E_{K_U}(M))=E_{K_U}(D_{K_R}(M))=M$。

公开密钥密码从根本上克服了传统密码在密钥分配上的困难。利用公开密钥密码进行保密通信需要成立一个密钥管理机构（KMC），每个用户都将自己的姓名、地址和公开的加密密钥等信息在 KMC 上注册登记，将公钥计入共享的公开密钥数据库 PKDB 中，KMC 负责密钥的管理，并且得到用户的信赖。这样，用户利用公开密钥密码进行保密通信就像查电话号码本打电话一样方便，无须按约定持有相同的密钥，因此特别适合计算机网络应用。

2.3.1　RSA

RSA 公钥密码算法是由美国麻省理工学院（MIT）的 Rivest，Shamir 和 Adleman 在 1978 年提出的，其算法的数学基础是初等数论的 Euler 定理，其安全性建立在大整数因子分解的困难性之上。

RSA 密码体制的明文空间 $M=$ 密文空间 $C=Zn$ 整数，其算法描述如下。

① 密钥的生成：首先，选择两个互异的大素数 p 和 q（保密），计算 $n=pq$（公开）$\varphi(n)=(p-1)(q-1)$（保密），选择一个随机整数 $e(0<e<\varphi(n))$，满足 $\gcd(e,\varphi(n))=1$（公开）。计算 $d=\mathrm{e}^{-1}\mathrm{mod}\varphi(n)$（保密）。确定公钥 $K_e=\{e,n\}$，私钥 $K_d=\{d,p,q\}$，即$\{d,1\}$。

② 加密：$C=M^e \bmod n$

③ 解密：$M=C^d \bmod n$

例如，$p=17$，$q=11$，$e=7$，$M=88$，使用 RSA 算法计算密文 C：

① 选择素数 $p=17$，$q=11$；

② 计算 $n=pq=187$；

③ 计算 $\varphi(n)=(p-1)(q-1)=160$；

④ 选择 $e=7$，满足 $0<e<160$，且 $\gcd(7,160)=1$；

⑤ 计算 d，因为 $d=e^{-1}\bmod\varphi(n)$，即 $ed\equiv 1\bmod\varphi(n)$，选择 $d=23$，因为 $23\times 7=1\times 160+1$；

⑥ 公钥 $K_e=\{e,n\}=\{7,187\}$，私钥 $K_d=\{d,n\}=\{23,187\}$；

⑦ 计算密文 $C=M^e \bmod n=88^7 \bmod 187=11$。（解密 $M=11^{23} \bmod 187=88$）

由于 RSA 密码安全、易懂，既可用于加密，又可用作数字签名，因此 RSA 方案是唯一被广泛接受并实现的通用公开密钥密码算法，许多国家标准化组织，如 ISO、ITU 和 SWIFT 等都已接受 RSA 作为标准。Internet 网的 E－mail 保密系统 PGP（Pretty Good Privacy）及国际 VISA 和 MASTER 组织的电子商务协议（Secure Electronic Transaction，SET 协议）中都将 RSA 密码作为传送会话密钥和数字签名的标准。

2.3.2　ElGamal 和 ECC

ElGamal 密码是除了 RSA 密码之外最有代表性的公开密钥密码。ElGamal 密码建立在离散对数的困难性之上。由于离散对数问题具有较好的单向性，所以离散对数问题在公钥密码学中得到了广泛应用。除了 ElGamal 密码外，Diffie－Hellman 密钥分配协议和美国数字签名标准算法 DSA 等也都是建立在离散对数问题之上的。ElGamal 密码改进了 Diffie 和 Hellman 的基于离散对数的密钥分配协议，提出了基于离散对数的公开密钥密码和数字签名体制。由于 ElGamal 密码的安全性建立在 GF(p)离散对数的困难性之上，而目前尚无求解 GF(p)离散对数的有效算法，所以 p 足够大时 ElGamal 密码是很安全的。

椭圆曲线密码体制（Elliptic Curve Cryptography，ECC）通过“元素”和“组合规则”组成群的构造方式，使得群上的离散对数密码较 RSA 密码体制而言能更好地对抗密钥长度攻击，使用椭圆曲线公钥密码的身份加密系统能够较好地抵御攻击，是基于身份加密的公钥密码学在理论上较为成熟的体现。由于椭圆曲线密码学较难，我们在这里不详细介绍。

2.3.3　公钥密码体制应用

大体上说，可以将公开密钥密码系统的应用分为如下三类。

1. 机密性的实现

发送方用接收方的公开密钥加密报文，接收方用自己相应的私钥来解密，如图 2-4 所示。

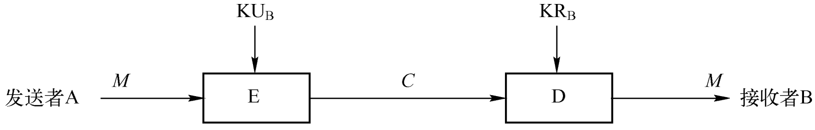


图 2-4　公开密钥算法加密过程

发送者 A 发送的信息用接收者 B 的公钥 KU_B进行加密，只有拥有与公钥匹配的私钥 KR_B的接收者 B 才能对加密的信息进行解密，而其他攻击者由于并不知道 KR_B，因此不能对加密信息进行有效解密。此加密过程保证了信息传输的机密性。

2. 数字签名

数字签名是证明发送者身份的信息安全技术。在公开密钥加密算法中，发送方用自己的私钥“签署”报文（即用自己的私钥加密），接收方用发送方配对的公开密钥来解密以实现认证，如图 2-5 所示。

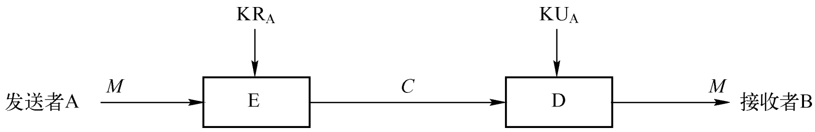


图 2-5　公开密钥算法数字签名过程

发送者 A 用自己的私钥 KR_A对信息进行加密（即签名），接收者用与 KR_A匹配的公钥 KU_A进行解密（即验证）。因为只有 KU_A才能对 KR_A进行解密，而发送者 A 是 KR_A的唯一拥有者，因此可以断定 A 是信息的唯一发送者。此过程保证了信息的不可否认性。

3. 密钥交换

密钥交换即发送方和接收方基于公钥密码系统交换会话密钥。这种应用也称混合密码系统，可以通过常规密码体制加密需要保密传输的消息本身，然后用公钥密码体制加密常规密码体制中使用的会话密钥，充分利用了对称密码体制在处理速度上的优势和非对称密码体制在密钥分发和管理方面的优势，从而使效率大大提高，如图 2-6 所示。

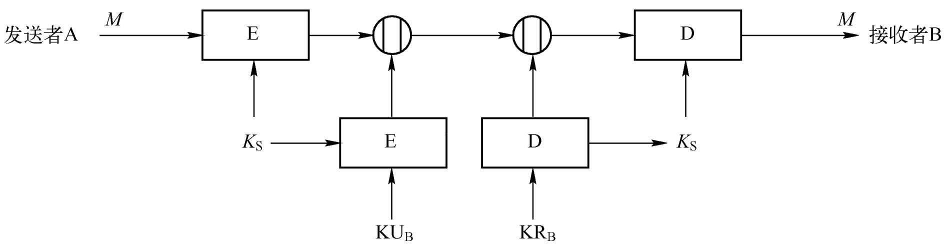


图 2-6　公开密钥算法交换会话密钥过程

发送者 A 发送明文用常规加密算法进行加密，然后把会话密钥 K_S用接收者的公钥 KU_B加密并与密文一起发送出去。接收者接收到信息后先对信息进行分类，将加密密钥用自己的私钥 KR_B进行解密得到会话密钥 K_S，然后再用 K_S对密文进行解密恢复出明文。

2.4　认证与数字签名

2.4.1　Hash 函数

1. Hash 函数的特性

Hash 函数也被称为哈希函数或散列函数。它是一种单向密码体制，即它是一个从明文

到密文的不可逆映射，能够将任意长度的消息 M 转换成固定长度的输出 $H(M)$。

Hash 函数除了上述特点之外，还必须满足以下三个性质：

① 给定 M，计算 $H(M)$ 是容易的；

② 给定 $H(M)$，计算 M 是困难的；

③ 给定 M，要找到不同的消息 M'，使得 $H(M)=H(M')$ 是困难的。实际上，只要 M 和 M' 略有差别，它们的散列值就会有很大不同，而且即使修改 M 中的一比特，也会使输出的比特串中大约一半的比特发生变化，即具有雪崩效应。注意：不同的两个消息 M 和 M' 使得 $H(M)=H(M')$ 是存在的，即发生了碰撞，但按要求找到一个碰撞是困难的，因此 Hash 函数仍可以放心地使用。

2. Hash 函数的算法

单向散列函数的算法有很多种，如 Snefru 算法、N－Hash 算法、MD2 算法、MD4 算法、MD5 算法等，常用的有 MD5 算法和 SHA－1 算法。

① MD5 算法：MD 表示信息摘要（Message Digest）。MD4 是 Ron Rivest 设计的单向散列算法，其公布后由于有人分析出算法的前两轮存在差分密码攻击的可能，因而 Rivest 对其进行了修改，产生了 MD5 算法。MD5 算法将输入文本划分成 512 位的分组，每一个分组又划分为 16 个 32 位的子分组，输出由 4 个 32 位的分组级联成一个 128 位的散列值。

② 安全散列算法（SHA）由美国国家标准和技术协会（NIST）提出，在 1993 年公布并作为联邦信息处理标准（FIPS PUB 180），之后在 1995 年发布了修订版 FIPS PUB 180，通常称之为 SHA－1。SHA 是基于 MD4 算法的，在设计上很大程度是模仿了 MD4。SHA－1 算法将输入长度最大不超过 264 位的报文划分成 512 位的分组，产生一个 160 位的输出。

由于 MD5 和 SHA－1 都是由 MD4 导出的，因此两者在算法、强度和其他特性上都很相似。它们之间最显著和最重要的区别在于 SHA－1 的输出值比 MD5 的输出值长 32 位，因此 SHA－1 对强行攻击有更强大的抵抗能力。MD5 算法的公开，使它的设计容易受到密码分析的攻击，而有关 SHA－1 的标准几乎没有公开过，因此很难判定它的强度。另外，在相同硬件条件下，由于 SHA－1 运算步骤多且要求处理 160 位的缓存，因此比 MD5 仅处理 128 位缓存的速度要慢。SHA－1 与 MD5 两个算法的共同点是算法描述简单、易于实现，并且不需要冗长的程序或很大的替代表。

2.4.2 报文认证

报文认证是证实收到的报文来自可信的源点并未被篡改的过程。常用的报文认证函数包括报文加密、散列函数和报文认证码三种类型。

1. 报文加密

报文加密是用整个报文的密文作为报文的认证符。发送者 A 唯一拥有密钥 K，如果密文被正确恢复，则 B 可以知道收到的内容没有经过任何改动，因为不知道 K 的第三方想要根据所期望的明文来找出能够被 B 恢复的密文是非常困难的。因此对报文进行加密既能保证报文的机密性，又能认证报文的完整性。报文加密认证过程如图 2-7 所示。

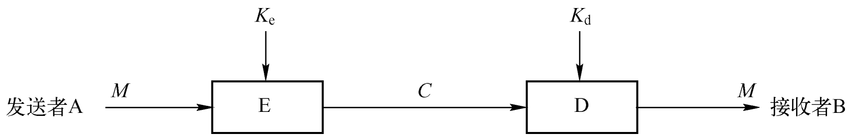


图 2-7　报文加密认证过程

2. 散列函数

散列函数是一个将任意长度的报文映射为定长的散列值的公共函数，并以散列值作为认证码。发送者首先计算要发送的报文 M 的散列函数值 $H(M)$，然后将其与报文一起发给 B，接收者对收到的报文 M' 计算新的散列函数值 $H(M')$ 并与收到的 $H(M)$ 进行比较，如果两者相同则证明信息在传送过程中没有遭到篡改。用散列函数进行认证的过程如图 2-8 所示。

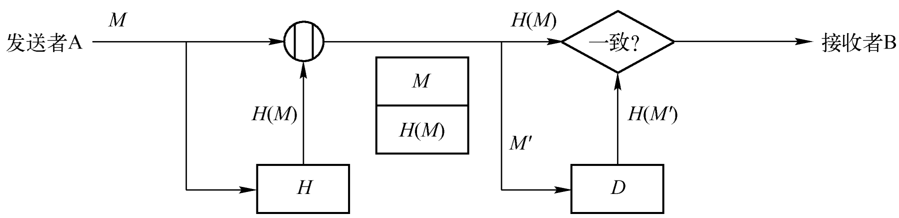


图 2-8　散列函数认证过程

3. 报文认证码

报文认证码（Message Authentication Code，MAC）是一个报文的公共函数和用于产生一个定长值的密钥的认证符。它使用一个密钥产生一个短小的定长数据分组，即报文认证码 MAC，并把它附加在报文中。发送者 A 用明文 M 和密钥 K 计算要发送报文的函数值 C_K（M），即 MAC 值，并将其与报文一起发送给 B，接收者用收到的报文 M 和与 A 共有的密钥 K 计算新的 MAC 值并与收到的 MAC 值进行比较。如果两者相同则证明信息在传送过程中没有遇到篡改。用 MAC 进行认证的过程如图 2-9 所示。

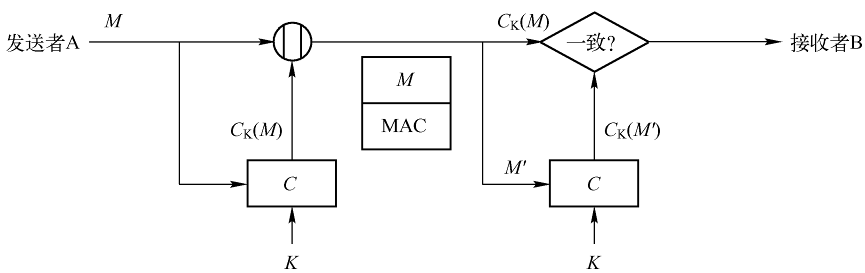


图 2-9　MAC 认证过程

基于对称分组密码（如 DES）是构建 MAC 最常用的方法，但由于散列函数（如 MD5 和 SHA-1）的软件执行速度比分组密码快、库函数容易获得及受美国等国家出口限制等原因，MAC 的构建逐步转向由散列函数导出。由于散列函数（如 MD5）并不是专门为 MAC 设计的，不能直接用于产生 MAC，因此提出了将一个密钥与现有散列函数结合起来的算法，

其中最具有代表性的是 HMAC（RFC2104）。HMAC 已经作为 IP 安全中强制执行的 MAC，并且也被 SSL 等其他的 Internet 协议所使用。

2.4.3 数字签名

数字签名与手写签名一样，不仅要能证明消息发送者的身份，还要能与发送的信息相关。它必须能证实作者身份和签名的日期和时间，必须能对报文内容进行认证，并且还必须能被第三方证实以便解决争端。其实质就是签名者用自己独有的密码信息对报文进行处理，接收方能够认定发送者唯一的身份，如果双方对身份认证有争议则可由第三方（仲裁机构）根据报文的签名来裁决报文是否确实由发送方发出，以保证信息的不可抵赖性，而对报文的内容及签名的时间和日期进行认证是为了防止数字签名被伪造和重用。

常用的数字签名采用公开密钥加密算法来实现，如采用 RSA、ElGamal 签名来实现。在图 2-5 中已经演示了发送者用自己的私钥对报文进行签名，接收者用发送者的公钥进行认证的过程。但由于直接用私钥对报文进行加密不能保证信息的完整性，因此，必须和散列函数结合起来实现真正实用的数字签名，如图 2-10 所示。

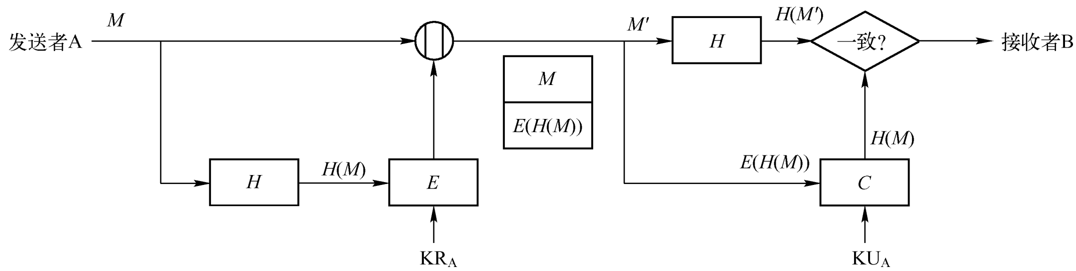


图 2-10 公开密钥算法数字签名

发送者用自己的私钥对信息的 Hash 值进行加密，然后与明文进行拼接发送出去。接收者一方面对收到的明文信息重新计算 Hash 值，一方面对前面的信息用发送者的公钥进行验证，得到的 Hash 值与重新计算的 Hash 值进行比较，如果一致，则说明信息没有被篡改。这种方法的优点在于在保证发送者真实身份的同时，还保证了信息的完整性，满足了数字签名的要求；不足之处是由于数字签名并不对明文进行处理，因此不能保证消息的机密性，但可以对信息进行加密，接收方收到信息后用自己的私钥进行解密，再验证数字签名及信息的完整性，如图 2-11 所示。

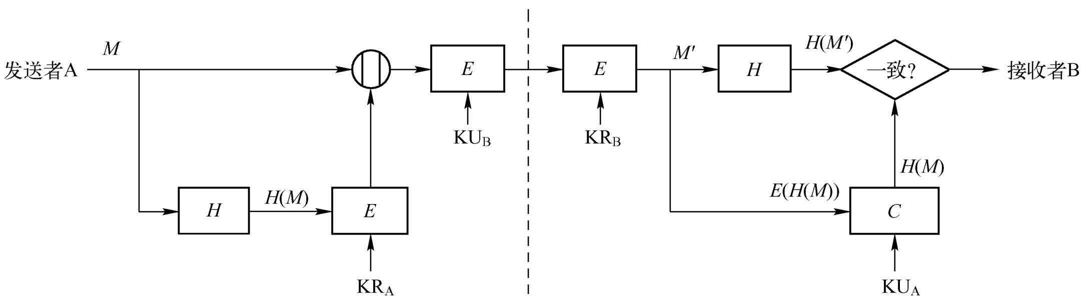


图 2-11 带数字签名信息的秘密通信

数字签名的另一种算法是使用仲裁机构进行签名，如采用常规加密算法与仲裁机构相结合实现数字签名。假设发送者 A 与接收者 B 用密钥 K_{AB}进行通信，仲裁者为 C，发送者 A 与仲裁者 C 之间共享密钥 K_{AC}，接收者 B 与仲裁者 C 之间共享密钥 K_{BC}，签名过程如图 2-12所示。

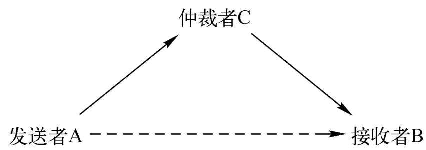


图 2-12　常规加密算法与仲裁机构相结合的数字签名

采用第三方仲裁机构进行数字签名的方法归纳见表 2-1。

表 2-1　采用第三方仲裁机构进行数字签名的方法

	发送者 A -> 仲裁者 C	仲裁者 C -> 接收者 B	数字签名 $S_A(M)$说明
常规加密：仲裁者 C 能看见明文	$M \parallel E_{K_{AC}}(S_A(M))$	$E_{K_{BC}}(ID_A \parallel M \parallel E_{K_{AC}}(S_A(M)) \parallel T)$	$S_A(M)=ID_A \parallel H(M)$
常规加密：仲裁者 C 不能看见明文	$ID_A \parallel E_{K_{AB}}(M) \parallel E_{K_{AC}}(S_A(M))$	$E_{K_{BC}}(ID_A \parallel E_{K_{AB}}(M) \parallel E_{K_{AC}}(S_A(M)) \parallel T)$	$S_A(M)=ID_A \parallel H(E_{K_{AB}}(M))$
公开密钥加密：仲裁者 C 不能看见明文	$ID_A \parallel E_{KR_A}(ID_A \parallel E_{KU_B}(S_A(M))$	$ID_A \parallel E_{KR_C}(ID_A \parallel E_{KU_B}(S_A(M)) \parallel T$	$S_A(M)=E_{KR_A}(M)$

① 常规加密且仲裁者 C 能看见明文：发送者 A 将发送的明文和签名信息 $S_A(M)$用密钥 K_{AC}加密后发送给仲裁者 C，C 对信息进行解密，恢复出明文 M，对 A 的签名信息$S_A(M)$进行认证，确认正确后将明文信息和签名信息及时间戳 T 用接收者 B 共享的密钥 K_{BC}加密后，发送给接收者 B。由于接收者 B 完全信任仲裁者 C，因此可以确信它发过来的信息就是发送者 A 发的信息。其中发送者 A 的签名信息 $S_A(M)$由 A 的标识符 ID_A和明文的散列函数值 $H(M)$组成。

② 常规加密且仲裁者 C 不能看见明文：如果不想被仲裁者 C 看见明文，则可用发送者 A 和接收者 B 之间的共享密钥 K_{AB}对明文进行加密，与签名一起发出。仲裁者 C 只能对发送者 A 的签名进行认证，但不能解密密文。只有接收者 B 能够对仲裁者 C 转发的信息进行两次解密，得到明文。

③ 公开密钥加密且仲裁者 C 不能看见明文：发送者 A 用自己的私钥 KR_A对信息进行签名，然后用接收者 B 的公钥 KU_B进行加密，再用私钥 KR_A对所有信息进行签名。仲裁者 C 收到信息后用发送者 A 的公钥 KU_A进行认证，然后用自己的私钥 KR_C对信息进行签名并转发给接收者 B。接收者 B 收到信息后通过仲裁者 C 的公钥 KU_C进行签名认证，确认发送者 A 的密钥是有效的，再用 A 的公钥 KU_A对信息进行解密恢复出明文。

2.5　密钥管理与分发

密钥管理负责密钥从产生到最终销毁的整个过程，包括密钥的生成、存储、分配、使用、备份恢复、更新、撤销和销毁等，是提供机密性、完整性和数字签名等密码安全技术的基础。

密钥的分发是保密通信中的一方生成并选择密钥，然后把该密钥发送给参与通信的其他一方或多方的机制，一般分为秘密密钥分发和公开密钥分发两大类。

1. 秘密密钥分发

秘密密钥分发使用一个大家都信任的密钥分发中心（Key Distrubution Center，KDC），每一通信方与 KDC 共享一个密钥，其交换方式有如下两种。

① 秘密密钥交换方式一，如图 2-13 所示。

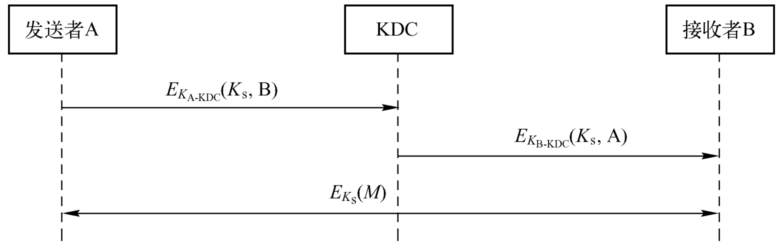


图 2-13　秘密密钥交换方式之一

- 发送者 A 随机生成会话密钥 K_S，然后将 K_s 和要通信的对象信息 B 用 A 与 KDC 之间的共享密钥 K_{A-KDC}加密后发送给 KDC。
- KDC 将收到的信息解密后得到会话密钥 K_S，将此密钥和信息发送者 A 的身份信息用 B 和 KDC 之间共享的密钥 K_{B-KDC}加密后发送给接收者 B。
- B 收到加密信息后进行解密，得到用 K_S 与 A 通信的信息。
- 发送者 A 与接收者 B 之间用会话密钥 K_S 进行信息的秘密传送。

② 秘密密钥交换方式二，如图 2-14 所示。

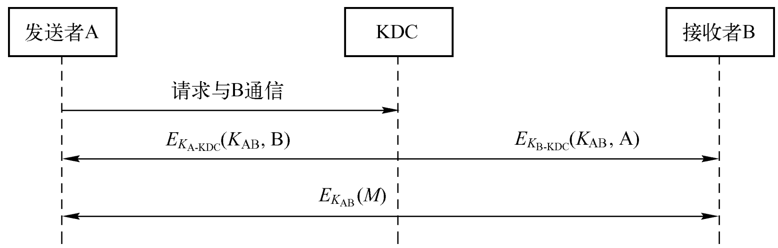


图 2-14　秘密密钥交换方式之二

- 发送者 A 将要与 B 之间进行通话的请求发送给 KDC。
- KDC 随机生成会话密钥 K_{AB}，并将 B 的身份信息用 A 与 KDC 之间的共享密钥 K_{A-KDC}加密后一起发送给 A，同时将会话密钥 K_{AB}与 B 的身份信息用 B 与 KDC 之间的共享密钥 K_{B-KDC}加密后发送给 B。
- 发送者 A 与接受者 B 分别对收到的 KDC 加密信息进行解密，得到通信另一方的信息和会话密钥 K_{AB}。
- 发送者 A 与接收者 B 用会话密钥 K_{AB}进行信息的加密传输。

2. 公开密钥分发

公开密钥分发方法有以下四种。

① 直接将密钥发送给通信的另一方或通过广播的方式将公钥发送给通信的其他方。

② 建立一个可动态访问的公钥目录（存放公钥信息的数据库服务器），使通信的各方可以基于公开渠道访问公钥目录来获取密钥。

③ 带认证功能的在线服务器公钥分发。例如，发送者 A 向管理员请求接收者 B 的公钥，管理员将接收者 B 的公钥用自己的私钥进行签名，使发送者 A 能够通过管理员的公钥进行认证，从而确认接收者 B 的公钥是可以用的。

④ 使用数字证书进行公钥分发。在这个方案中，一般有一个可信的第三方机构——认证中心，又称证书授权（Certificate Authority，CA），通过各方向 CA 申请证书并信任 CA 颁发的证书。证书的内容一般包括证书的 ID、证书的发放者、证书的有效期、用户的名称、用户的公钥，以及证书发放者对证书内容的签名信息。用户能够根据 CA 的签名信息来认证证书的有效性和合法性，并利用证书中的用户公钥对信息进行加密通信。CA 负责数字证书的颁发和管理。

参考文献

[1] 杨寅春，等. 网络安全技术 [M]. 西安：西安电子科技大学出版社，2009.
[2] 胡向东，魏琴芳. 应用密码学教程 [M]. 北京：电子工业出版社，2005.
[3] 邓亚平. 计算机网络安全 [M]. 北京：人民邮电出版社，2004.

第 3 章　感知层物理安全技术

物联网感知层设备主要用于信息采集和目标检测等领域，通常部署在极端的网络环境中，如水下、战场、野外等，使得设备的管理和维护都非常困难，一旦有感知层设备被攻击者捕获并破解，则管理人员很难发现，因此需要增强感知层设备自身的物理安全防护技术。本章在对 RFID、智能卡、传感器网络节点面临的物理安全威胁进行介绍的基础上，重点给出了 RFID、智能卡、传感器节点的物理安全防护技术；同时针对感知层设备面临的克隆攻击问题，介绍了物理不可复制功能 PUF 及基于 PUF 的抗克隆攻击技术。

3.1　RFID 标签物理层安全威胁及防护技术

RFID（Radio Frequency Identification，射频识别）是一种利用射频信号自动识别目标对象并获取相关信息的技术。RFID 最早的应用可追溯到第二次世界大战中用于区分盟军和纳粹飞机的“敌我辨识”系统。随着技术的进步，RFID 应用领域日益扩大，现已涉及日常生活的各个方面，并将成为未来信息社会建设的一项基础技术。RFID 典型应用包括：在物流领域用于仓库管理、生产线自动化、日用品销售，在交通运输领域用于集装箱与包裹管理、高速公路收费与停车场收费，在农牧渔业用于羊群、鱼类、水果等的管理及宠物、野生动物跟踪，在医疗行业用于药品生产、病人看护、医疗垃圾跟踪，在制造业用于零部件与库存的可视化管理。RFID 还可以应用于图书与文档管理、门禁管理、定位与物体跟踪、环境感知和支票防伪等多种应用领域。在涉密信息系统方面，RFID 技术可以用于涉密信息设备的管理。此外，电子标签也可以用于对各种形式的介质存储的涉密文件进行管理。

3.1.1　RFID 标签的破解及复制

一个典型的 RFID 应用系统通常包括以下部分。

① 后台系统：部署在一台或多台服务器上的应用软件和数据库。

② 读写器系统：读写器负责与电子标签的通信，这些非接触式的通信需要通过读写器内置的或外界的填写来完成，对于无源电子标签，读写器系统还要通过天线为标签提供能量。

③ 标签：标签内集成了天线和芯片，芯片内存储了标签 ID 并具有较为有限的计算能力，一些芯片内还会存储其他信息，芯片按照其内容是否可写分为只读标签和可写标签。

在 RFID 安全相关的研究中，通常假设从后台系统到读写器系统的信道是安全的，而从读写器系统到电子标签的信道是不安全的，如图 3-1 所示。

如图 3-2 所示，一个 RFID 系统的信道可以分为以下三类。

① 前向信道（Forward Channel）：前向信道承载从读写器到标签的通信。

② 后向信道（Backward Channel）：后向信道承载从标签到读写器的通信。

③ 内存信道（Memory Channel）：内存信道承载标签内部的信息传输。

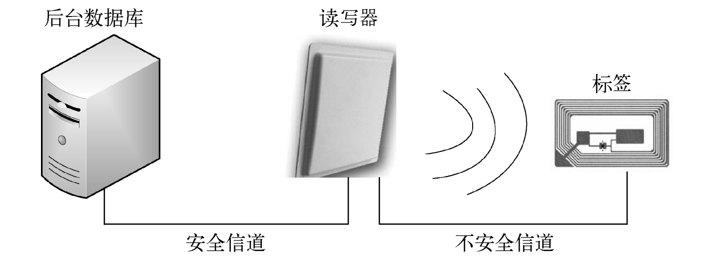


图 3-1　RFID 系统的基本假设

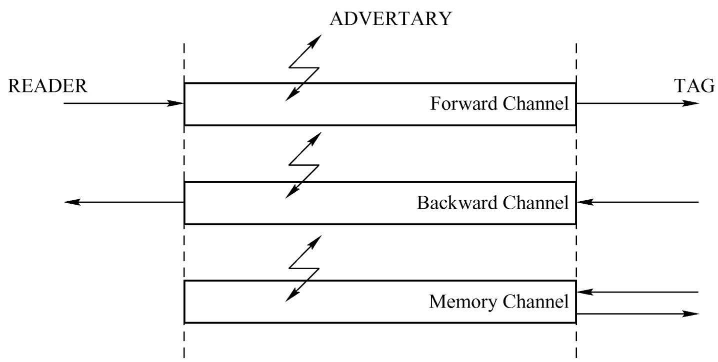


图 3-2　RFID 系统信道的分类

通常情况下，前向信道的通信距离要大于后向信道的通信距离，在使用无源电子标签时，这种现象尤为显著。

RFID 系统包括标签、读写器及标签与读写器之间的射频通信信道。在使用电子标签进行交易的业务中，标签复制和伪造会给使用者带来损失；在 RFID 标签应用较广的供应链中，如何防止信息的窃听和篡改显得尤为重要。

RFID 标签的安全问题主要包括以下几方面。

1. 信息传输安全问题

物联网终端很多时候都通过无线电波传输信号，智能物品感知信息和传递信息基本上都是通过无线传输实现的，这些无线信号存在被窃取、监听和其他危险。目前，在信息传输中攻击者使用的主要方式可分为两大类，即主动攻击和被动攻击，主动攻击中最常见的攻击手段为信道堵塞，而被动攻击主要以监听和窃听技术为主。

2. 数据真实性问题

电子标签的身份识别在物联网系统中非常重要。攻击者可以从窃听到的标签与读写器间的通信数据中获得敏感信息，进而重构 RFID 标签，达到伪造标签的目的。攻击者可以利用伪造标签替换原有的标签，或通过重写合法的 RFID 标签内容，使用低价物品的标签替换高价物品的标签从而非法获益。同时，攻击者也可以通过某种方式隐藏标签，使读写器无法发现该标签，从而成功地实施物品转移。读写器只有通过身份认证才能确信消息是从正确的标签发送过来的。

3. 信息和用户隐私泄露问题

信息泄露是指 RFID 标签发送的信息被暴露，该消息包括标签用户或识别对象的相关信息，这些信息一般包含一些用户的隐私和其他敏感数据。例如，RFID 物流商品通信信息是公开的，其他任何人都可以获得收发双方及物品的信息。当电子标签应用于药品时，很可能暴露药物使用者的病理，隐私侵犯者可以通过扫描服用的药物推断出某人的健康状况。一个安全的 RFID 系统必须拥有一个安全的 RFID 标签，从而保护用户的隐私信息或相关经济实体的商业利益。

4. 数据秘密性问题

安全的物联网方案应该可以保证标签中的信息只能被收取读写器识别。但是目前读写器和标签的通信是不受保护的，未采用安全机制的 RFID 标签会向邻近的读写器泄露标签内容和一些敏感信息。如果缺乏支持点对点加密和 PKI 密钥交换的功能，在物联网系统的应用过程中，攻击者能够获取并利用 RFID 标签上的内容。

5. 数据完整性问题

在通信过程中，数据完整性能够保证接收者收到的信息在传输过程中没有被攻击者篡改和替换。在基于公钥的密码体制中，数据完整性一般是通过数字签名完成的。在 RFID 系统中，通常使用消息认证码进行数据完整性的检验，它使用的是一种带有共享密钥的散列算法，即将共享密钥和带验证的消息连接在一起进行散列运算，对数据的任何细微改动都会对消息认证码的值产生较大的影响。事实上，除了采用 ISO14443 标准的系统（该系统使用了消息认证码）外，在读写器和标签的通信过程中，传输信息的完整性无法得到保障。在通信接口处使用校验和的方法也仅仅能够检测随机错误的发生。如果不采用数据完整性控制机制，可写的标签存储器有可能受到攻击。攻击者编写软件，利用计算机的通信接口，通过扫描 RFID 标签和响应读写器的查询，寻找安全协议、加密算法及其实现机制上的漏洞，进而删除或篡改 RFID 标签内的数据。

6. 恶意追踪

随着 RFID 技术的普及，标签识别装备的价格也越来越低廉，特别是 RFID 进入人们的日常生活后，拥有阅读器的人可以扫描并追踪别人，而且被动标签信号不能切断、尺寸很小，极易隐藏并且使用寿命很长，可以自动化识别和采集数据，这就加剧了恶意追踪的问题。

3.1.2　RFID 标签的物理安全防护技术

RFID 系统容易遭受各种主动和被动攻击的威胁，RFID 系统本身的安全问题可归纳为隐私和认证两个方面：在隐私方面主要是可追踪性问题，即如何防止攻击者对 RFID 标签进行任何形式的跟踪；在认证方面主要是要确保只有合法的阅读器才能够与标签进行交互通信。当前，在物理方法保障 RFID 系统本身安全的方面主要有 Kill 命令、静电屏蔽、主动干扰及阻塞标签。随着物联网技术的日益普及应用，使用 RFID 标签的消费者隐私权备受关注。

1. Kill 命令机制（Kill 标签）

Kill 命令机制是由标准化组织自动识别中心（Auto – ID Center）提出的。Kill 命令机制采用从物理上销毁 RFID 标签的方法，一旦对标签实施了销毁（Kill）命令，RFID 标签将永久作废。阅读器无法再对销毁后的标签进行查询和发布指令，通过自戕的方法保护了消费者的个人隐私。这种牺牲 RFID 标签功能及后续服务的方法可以在一定程度上阻止扫描和追踪。但是 Kill 命令机制的口令只有 8 位，因此恶意攻击者仅以 2^8 的计算代价就可以获得标签的访问权。而且由于电子标签销毁后不再有任何应答，很难检测是否真正对标签实施了 Kill 操作。因此，Kill 标签并非是一个有效检测和阻止标签扫描与追踪的防止隐私泄露技术。

2. 静电屏蔽机制

静电屏蔽机制的工作原理是使用法拉第网罩来屏蔽标签。如图 3–3 所示，法拉第网罩是由金属网或金属箔片构成的阻隔电磁号穿透的容器。添加法拉第网罩前（图 3–3 左边部分）两个物体可产生电磁反应，但加了法拉第网罩后（图 3–3 右边部分），外部电磁信号不能进入法拉第网罩，里面的磁波电信号也无法穿透出去。当人们把标签放进由传导材料构成的容器里时可以阻止标签被扫描，被动标签接收不到信号也就不能获得能量，主动标签发射的信号也不能发出。利用法拉第网罩可以阻止非法窥测者通过扫描获得标签的信息。采用法拉第网罩需要添加一个额外的物理设备，带来了不方便，也增加了物联网系统设备的成本。

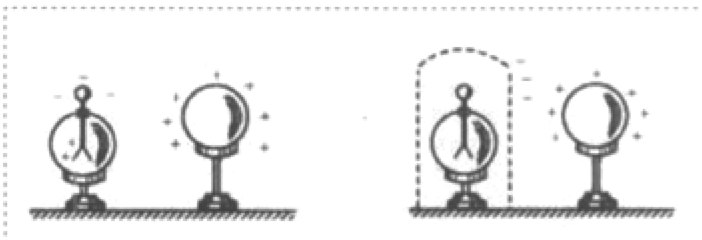


图 3–3　法拉第网罩

3. 主动干扰

主动干扰无线电信号是另一种屏蔽标签的方法。标签用户可以通过一个设备主动广播无线电信号以阻止或破坏附近的物联网阅读器的操作。这种初级的方法可能导致非法干扰。附近其他的合法的物联网系统也会受到干扰，更严重的是它可能阻断附近其他使用无线电信号的系统。

4. 阻塞标签法

阻塞标签（Blocker Tag）法通过阻止阅读器读取标签确保消费者隐私。与一般用来识别物品的标签不同，Blocker Tag 是一种被动干扰器。当阅读器在进行某种分离操作时，当搜索到 Blocker Tag 所保护的范围时，Blocker Tag 便发出干扰信号，使阅读器无法完成分离动作，阅读器无法确定标签是否存在，也就无法和标签沟通，由此来保护标签，保护用户的隐私。但是由于增加了阻塞标签，因此应用成本相应增加。其次，Blocker Tag 可以模拟大量的标签

ID，从而阻止阅读器访问隐私保护区域以外的其他标签，因此 Blocker Tag 的滥用可能导致拒绝服务攻击。同时，Blocker Tag 有其作用范围，超出隐私保护区域的标签将得不到保护。

5. RFID 标签芯片的攻击

对 RFID 标签的攻击包括破坏性攻击和非破坏性攻击，其中针对破坏性攻击主要有存储器读出技术和版图重构两种防范措施。

1）存储器读出技术

存放密钥、用户数据等内容的存储器不能通过简单的光学照片获得其中的信息。在安全认证过程中，至少要对这些数据区访问一次，因此，可以使用微探针监听总线上的信号获取重要数据。顶层探测器网格是有效防止微探针获取存储器数据的重要手段之一，充分利用深亚微米 CMOS 技术提供的多层金属，在重要的信号线顶层构成探测器网格，能够连续监测短路和断路。当有电时，它能防止激光切割或选择性的蚀刻去获取总线的内容。根据探测器输出，芯片可立即触发电路将非易失性存储器中的内容全部清零。这些网格对于其下的各层金属连线重构也有影响，因为蚀刻不是均匀的，上层金属的模式在下层可见，会给版图的自动重构带来很多麻烦。手动探针的目标尺寸一般在 1 微米左右，尖端小于 0.1 微米的探针台价格在几十万美元之上，且极难获得。一个精心设计的网格将使手动微探针攻击难以实施，一般的 FIB（Focus Ion Beam）修补技术也难以逾越。

2）版图重构

破坏性攻击的一个重要步骤是重构 RFID 芯片的版图。通过研究连接模式和跟踪金属连线穿越可见模块的边界，达到迅速识别芯片上的一些基本结构的目的，如数据线和地址线。对于 RFID 标签芯片的设计来说，射频模拟前端需要采用全定制方式实现，但是常采用 HDL 语言描述来实现包括认证算法在内的复杂控制逻辑，显然这种采用标准单元库综合的实现方法会加速设计过程，但是也给以反向工程为基础的破坏性攻击提供了极大的便利，这种以标准单元库为基础的设计可以使用计算机自动实现版图重构。因此，采用全定制的方法实现 RFID 的芯片版图会在一定程度上加大版图重构的难度。版图重构的技术也可用于获得只读型 ROM 的内容。ROM 的位模式存储在扩散层，用氢氟酸（HF）去除芯片各覆盖层后，根据扩散层的边缘就很容易辨认出 ROM 的内容。基于微处理器的 RFID 设计中，ROM 中可能不包含任何加密的密钥信息，但是它的确包含足够的 I/O、存取控制、加密程序等信息，这些在非破坏性攻击中尤为重要。因此，对于使用微处理器的 RFID 设计，推荐优先使用 Flash 或 EEPROM 等非易失性存储器存放程序。

非破坏性攻击主要针对具有微处理器的产品而言。微处理器本质上是成百上千个触发器、寄存器、锁存器和 SRAM 单元的集合，这些器件定义了处理器的当前状态，结合组合逻辑则可知道下一时钟的状态。常见的非破坏性攻击主要有电流分析攻击和故障攻击。

（1）电流分析攻击及防范措施

根据电流分析攻击实施的特点，可将其分为简单电源攻击（SPA）和差分电源攻击。原则上，RFID 的电源集成在 AFE（Analog Front End）的内部，似乎远离了电流分析的危险，然而实际上并非如此。通过在 RFID 天线和串联的分压电阻的两端直接加载符合规格的交流信号，RFID 负载反馈信号可以百倍于无线模式下的信号强度直接叠加在加载的交流信号上。

由于芯片的功耗变化与负载调制在本质上是相同的，因此，如果 AFE 的电源设计不恰当，则 RFID 微处理器执行不同内部处理的状态可能在串联电阻的两端交流信号上反馈出来。针对电流分析攻击的特点，芯片的功耗是一个重要的问题，就工作效率而言，串联方案的效率更高，更适合集成电路设计。但是就安全而言，并联方案是更理想的选择，因为通过并联泄放电路将电源幅度和纹波的变化控制在尽可能小的范围内，将电源电流消耗波动抑制在整流电路之后，这样天线两端的交流信号不能反映任何内部基带系统（主要是微处理器）状态的差异。

（2）故障攻击及防范措施

通过故障攻击可以导致一个或多个触发器位于病态，从而破坏传输到寄存器和存储器中的数据。在所知的 RFID 标签芯片非破坏性攻击中，故障攻击是实际应用中最有效的攻击技术之一。时钟故障和电源故障都是故障攻击的主要手段。通过简单地增加或降低时钟频率一个或多个半周期可以实施时钟故障攻击，这样会使部分触发器在合法的新状态到来之前就采样它们的输入。时钟故障攻击通常和电源故障结合在一起，在接触式 RFID 标签中通过组合时钟和电源波动，增加程序计数器内容而不影响处理器的其他状态。这样，RFID 标签的任意指令序列都可以被黑客执行，而程序员在软件编写中并没有什么很好的应对措施。RFID 标签为了有效抵御时钟故障攻击，除了采用时钟探测器外，更重要的是严格限制 RFID 设计的工作频率范围、载频的谐波品质因素、对称性的指标。潜在的故障技术仍需进一步探索，如通过将金属探针置于距离处理器几百个微米的高度，在几毫秒内施加几百伏特的电压，得到的电场强度足够改变附近的晶体管阈值电压。这些技术的应用价值和应对措施还有待进一步的研究。

3.2　传感器网络节点的物理安全威胁及其防卸技术

传感器网络是一个结构松散、开放的网络，常常需要部署在无人监控的环境中，因此传感器网络节点很容易受到节点破坏攻击和节点泄露攻击。节点破坏攻击是指物理破坏节点，使其不能工作，进而破坏网络。节点泄露攻击是指捕获节点后，通过各种方法攻击节点，获得节点内部代码等敏感信息，还可以克隆伪造节点并进入网络进行信息窃取等破坏活动。节点泄露攻击大多是针对嵌入式芯片进行的。物理安全是传感器网络安全的根本，如果节点的物理安全不能保证，则其他安全措施都是纸上谈兵。

3.2.1　节点破坏攻击及其防御

1. 盲物理破坏攻击及防御

盲物理破坏攻击是指攻击者使用爆炸物等破坏性设备攻击传感器网络节点部署的区域，破坏区域内的节点，使网络无法工作。例如在战场，攻击者通过发射大量炮弹破坏目标网络的部署区域。

针对盲物理破坏攻击，其防御手段在于研究如何在目标区域内合理部署传感器节点，使之能在被破坏的情况下尽量维持网络功能，提高网络的抗容错和容灾性能。

2. 捕获破坏攻击及防御

捕获破坏攻击是指攻击者通过探测器等工具在传感器网络部署区域寻找和破坏传感器节点。和盲物理破坏攻击不同，攻击者不希望破坏网络部署区域，因此捕获破坏攻击不使用炸药等爆炸物。

针对捕获破坏攻击，其防御的关键在于如何最大限度伪装节点，并设计良好的节点通信协议，一旦某一节点被俘，则及时通知其他网络节点，尽量降低被探测到的风险。

3.2.2　节点泄露攻击及其防御

1. JTAG 攻击及防御

很多微处理器都留有 JTAG 节点，如 MICA2 等传感器节点的 ATmegal128 芯片等。通过 JTAG 攻击是指攻击者使用编程器通过 JTAG 接口读取节点的程序和数据，可以在很短的时间内把 EEPROM、Flash、SRAM 中的信息读取出来，通过反汇编软件获得程序代码，从而进一步分析节点所存储的协议、密钥、数据等机密信息。攻击者还可以在修改节点代码后使其重新加入网络，充当间谍来进一步破坏网络。

为了防止未授权访问或拷贝单片机的内部程序，大部分单片机都带有加密锁定位或加密字节，以保护片内程序。如果在编程时机密锁定位被使能，就无法用普通编程器直接读取单片机内部的信息，这就是所谓的拷贝保护或锁定功能。然而，实际上这种保护措施非常脆弱，低档次、廉价的单片机为了节约成本很少提供这项功能。

2. 过错攻击及防御

过错产生攻击是指攻击者通过各种手段产生异常工作条件，如离子射线、电磁辐射、非正常电压等，使得芯片在运算过程中产生错误，通过分析错误获得有用信息。使用最多的过错产生攻击手段包括电压冲击和时钟冲击，低电压和高电压攻击可用来禁止保护电路工作或强制处理器执行错误操作，时钟瞬态跳变也许会复位保护电路而不会破坏受保护信息。

1）电压冲击攻击

电源和时钟的瞬态跳变可以在某些处理器中影响单条指令的解码和执行，微处理器要求在稳定的电压下工作，一个短而巧妙的脉冲可以引起单步的程序错误而微处理器仍能够继续执行程序。例如，突然出现的能量短脉冲可能会对存储的逻辑值产生改变，使得程序读出错误数据。许多加密算法都易受到这一类故障注入的影响，采用差分故障分析（DFA）技术将正确的与错误的密码编码相比较，从而分析出密钥。电压冲击攻击需要了解程序执行的流程，选择关键的时刻进行攻击。

2）时钟冲击攻击

每一个晶体管都可以模型化一个 RC 电路，有着固定的延迟，芯片内部每个路径的延迟也是固定的。单片机的最高工作频率由其最长路径决定。对于某工作频率为 5MHz 的单片机，如果提供 20MHz 的时钟，某些触发器的值就会发生异常，通过控制时钟频率，可以让

单片机执行各种完全不同于指令集的操作。

防范过错攻击的策略一方面是严格的电压、频率和温度检测，能够监测到部分攻击；另一方面可以通过软件防范，通过检查关键的程序流向及加密运算结果来实现故障监测。

3. 侵入攻击及防御

侵入攻击指攻击者物理剖析芯片，直接暴露芯片内部连线，然后观察、操控、干扰单片机以达到攻击目的。侵入攻击需要昂贵的设备，而且攻击周期比较长，需要专业的技术人员。侵入攻击一般先通过各种化学方法揭去芯片封装，然后寻找熔丝保护位并将其暴露在紫外线下一段时间，破坏保护位的保护作用，最后用编程器读出程序。一种防御方法是设置感知电路感知入侵，然后消耗数据。

4. 板级攻击及防御

板级攻击是指通过在电路板上监听信号线，或者切断、重连信号线来更改电路连接的方法获取数据、程序、密钥等信息。例如，攻击者可以在电路板上割断 Flash 与原处理器的连线，直接将其连接到攻击者的微处理器上，通过处理器读出 Flash 的内容，或者将恶意代码写入，然后恢复连接。因此最好不要使用外接存储器来存储机密信息，如果必须要用，则最好采用加密处理。

5. 旁路攻击及防御

旁路攻击是指攻击者通过观察电路中的某些物理量，如能量消耗、电磁辐射、时间等的变化规律，来分析节点的程序、密钥、数据等信息。目前应用最为广泛的旁路攻击方法是差分能量分析（Differential Power Attack，DPA）。芯片执行不同指令时执行不同的操作，消耗的电流也是不一样的，通过使用电子监测仪器测量芯片电流，通过统计技术分析执行的指令及处理的数据，攻击者可以获得芯片中的信息。传统的防御 DPA 的方法有两种：一种是随机数发生器产生额外的噪声和干扰信号；另一种是通过增加滤波电路来消除噪声。

6. 逻辑攻击及防御

逻辑攻击对嵌入式系统来说威胁很大，它利用软件、加密算法或安全协议的弱点和漏洞进行攻击，如缓冲区溢出攻击等。逻辑攻击取得成功的一个典型事件是对早期 ATMELA、AT89C 系列单片机的攻击。攻击者利用了该系列单片机擦除操作时序设计上的漏洞，使用自编程序在擦除加密锁定位后，停止下一步擦除芯片内程序存储器的数据，从而使加密的单片机变成没有加密的单片机，然后利用编程器读出片内程序。

3.2.3　传感器节点安全设计

传感器节点安全是其网络安全的基础，节点安全包括节点本身的物理安全和节点里的数据、程序安全，针对节点的攻击包括物理破坏节点和窃取节点程序、数据、密钥等信息。传感器节点物理安全设计可以从以下方面来考虑：

- 程序烧写后锁定 JTAG；
- 选用比较生僻的单片机来增大破解难度，也可选用带有安全存储功能的芯片；
- 尽量不要使用外接存储器，如果使用，尽量加密数据；
- 加入移动感知或其他检测机制，及时发现非法操作行为。

3.3 物理不可克隆函数（PUF）技术

3.3.1 PUF 概述

在无线传感器网络中，包括公钥签名技术在内的很多身份认证方法常常被用作抵抗网络攻击的重要手段。然而由于节点间身份认证中所需的信息往往都存储在节点内部，因此当某一节点被攻击者捕获后，攻击者很容易便可实现复制攻击，也就是说，当节点被攻击者捕获并复制后，其他节点基本无法对其身份进行有效识别。尤其是在移动传感器网络、延迟容忍传感器网络及 RFID 等系统中，由于节点或标签常常疏于维护，因此能否抵抗克隆攻击成为网络的一项重要需求。

麻省理工学院（MIT）的 Srini Devadas 教授和他的团队于 2005 年在硅谷成立了 Verayo 公司，并为世界首创了“芯片 DNA”验证方法。这种方法就是 PUF（Physical Unclonable Functions）技术。

物理不可克隆函数是指对一个物理实体输入一个激励，利用其不可避免的内在物理构造的随机差异输出一个不可预测的响应，这样一个抗物理克隆攻击的函数如图 3-4 所示。显然，它最主要的优势是可以抵抗物理克隆攻击的发生，同时相对于资源有限的物理实体来说，它的优势还包括这种不能被克隆的激励响应行为不仅可以实现一些与传统公钥密码一样的功能，而且还能大大减少计算、存储和通信开销。

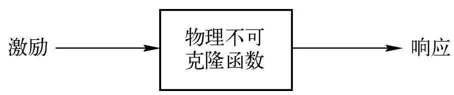


图 3-4　物理不可克隆函数的激励响应机制

从信息技术进化的角度，物理不可克隆函数的使用至少还有两个优势：首先，计算和通信设备往往变得更小，并且其深度整合会导致很多额外的物理相互影响，在这种情况下，自然地利用设备的物理性质，而不是刻意地消除这种影响会给系统减少额外的开销；第二，分布式（例如云）计算的日益发展和“物联网”的兴起使得数十亿的对象或计算设备可以互连，但是同时也造成了信任与安全的问题，在这种情况下，给每个对象或计算设备装备一个唯一的身份，可以作为更高水平安全架构的信托锚。

2001 年，Pappu 首先正式提出了物理不可克隆函数的概念，并设计实现光学 PUF 来实现系统认证等应用。从此，朝着实现方法和应用多样性的方向，人们提出了越来越多的物理不可克隆函数实现方法，如涂层 PUF、基于仲裁器的 PUF 和蝴蝶 PUF 等，并且基于这些实现方法实现越来越多新的安全应用，如知识产权（Intellectual Property，IP）保护、系统认证和密钥生成等。

3.3.2　PUF 基本原理及其数学模型

物理不可克隆函数的最基本原理是实现一个简单的验证机制，如图 3-5 所示。从图中可以看出，在注册阶段，后端服务器通过安全的通道把认证设备的激励响应对存储到数据库中；在认证阶段，在数据库中选择一个激励响应对，通过不可信的通道传递激励给需要认证的设备，设备通过 PUF 得到响应返回给数据库，如果两个响应一致，说明认证成功，否则认证失败。目前看来，虽然这样一个认证过程是不安全的，但是可以作为一个基本单元嵌入更复杂的协议之中来实现相应的功能。

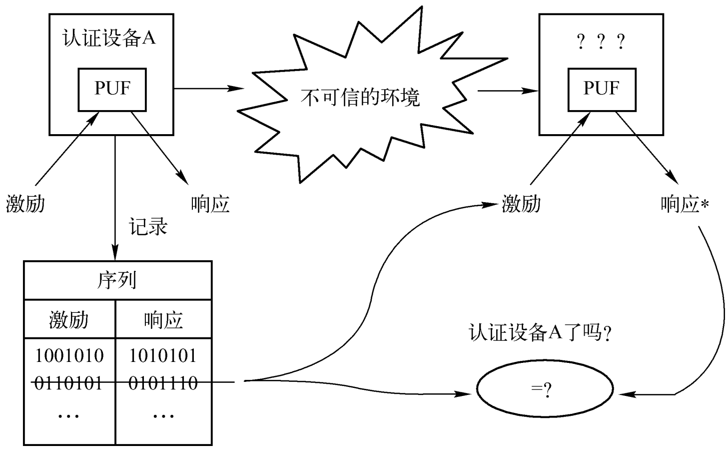


图 3-5　物理不可克隆函数的基本原理

Gassend 等人在“Controlled Physical Random Functions”一文中给出的 PUF 定义如下：

定义 3.1：A Physical Unclonable Function is a function that maps challenges to responses and that is embodied in a physical object. It satisfies the following properties：

1）Easy to evaluate：the physical object can be evaluated in a short amount of time.

2）Hard to characterize：from a number of measurements performed in polynomial time，an attacker who no longer has the device and who only has a limited（polynomial）amount of resources can only obtain a negligible amount of knowledge about the response to a challenge that is chosen uniformly at random.

从 PUF 的定义中可以看出，物理不可克隆函数是一个能够对输入的挑战产生响应的物理单元，且该单元与设备密不可分。PUF 具有响应快速、安全的特点，攻击者无法在有限的资源和时间的情况下获取 PUF 的有用信息。PUF 的数学模型介绍如下：假设在初始化阶段，挑战 C 输入到一个合法 PUF 中，并获得对应的响应 $R \in \Re^n$，其中（C，R）被称为挑战-响应对。同时假设在验证阶段合法的 PUF 对挑战 C 的响应为 $X \in \Re^n$，而一个伪造的 PUF 的对应挑战 C 的响应为 Y；另外假设 $F_{R,X,Y}$为随机变量 R、X 和 Y 的联合概率分布函数。

定义 3.2：假设 δ，ε_a，$\varepsilon_e \geqslant 0$，$(\Re^n)^3$ 上的联合分布函数 $F_{R,X,Y}$被称为$(\delta,\varepsilon_a,\varepsilon_e)$可靠的，当它满足以下两个条件：

① $\text{Prob}(d(R,X)>\delta)\leqslant\varepsilon_a$；

② $\text{Prob}(d(Y,R)\leqslant\delta)\leqslant\varepsilon_e$。

以上两个概率是在 $F_{R,X,Y}$上的联合分布概率。

从定义 3.2 可以看出，当向同一个 PUF 多次输入相同的挑战时，其输出的各响应相差较大的概率非常低；相反，向不同的 PUF 输入相同的挑战，则不同 PUF 的响应相差较大的概率非常高。在实际中，如果对 PUF 电路进行相关的纠错和萃取等处理后，可以将 PUF 看成输入为 C，输出为 R 的函数 P，因此有 $R=P(C)$。

PUF 的电路结构相对简单，可以在非常小的开销条件下嵌入芯片中，因此将 PUF 作为 RFID 标签、传感器节点的安全模块来提高设备的安全性具有非常大的实际意义。

3.3.3　PUF 分类及实现

自从 Pappu 正式提出物理不可克隆函数的概念以来的十几年间，人们提出了许多基于各种实现技术的新类型的物理不可克隆函数实现方法。本节根据电路技术，从纷繁复杂的实现方法中归类总结出三类，即非电子 PUF、模拟电路 PUF 和数字电路 PUF。目前提出的大部分物理不可克隆函数的实现方法可基本上归为以下几类，如图 3-6 所示。

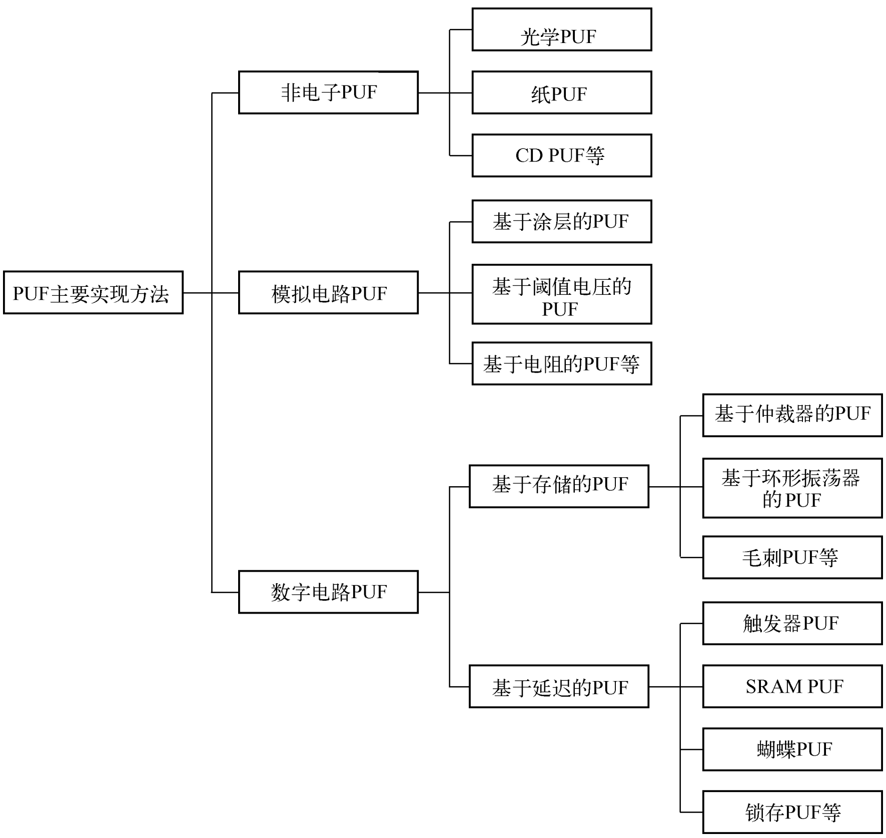


图 3-6　目前提出的主要物理不可克隆函数实现方法

1. 非电子 PUF

第一种物理不可克隆函数的实现方法是非电子 PUF，其中最具代表性的是 Pappu 提出的光学 PUF，如图 3-7 所示，它也是在安全应用中对物理不可克隆函数概念的第一个正

式描述。光学 PUF 的核心组件是一个随机掺杂光散射粒子的小光透明令牌，当用一个激光照射光透明令牌时，它就能辐射出一个有明暗斑点的复杂图像，即所谓的散斑。一个贾柏滤波器可以很好地提取这样一个散斑，并输出光学 PUF 的一个响应。因此，在光学 PUF 中，激光的物理参数（位置、方向、波长等）是激励，而滤波器的输出是响应。由于激光和散射粒子相互作用的复杂性质，响应被认为是高度随机和唯一的。光学 PUF 的响应依赖于光令牌的微观物理细节，这就会导致两个同样产生的令牌将显示出一个根本不同的激励响应行为，从而防止克隆的发生。此外，Pappu 提出令牌的一个小的物理变化，如钻微观孔等，将会大大改变物理不可克隆函数的激励响应行为，也就是说，其具有防篡改的属性。其他的非电子 PUF 包括纸 PUF 的概念，它的主要原理是利用纸文件不规则的纤维结构的激光反射来作为防止伪造的“指纹”。也有 CD PUF 的概念，它的主要原理是在 CD 制造过程中，可变因素会影响平坦面的精确长度和光盘的坑，这可以被用来提取 CD 的“指纹”。

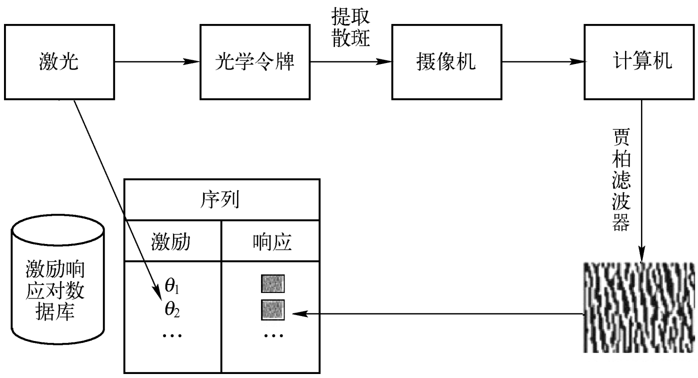


图 3-7　光学 PUF

2. 模拟电路 PUF

第二种物理不可克隆函数的实现方法是模拟电路 PUF，其中最具代表性的是涂层 PUF，如图 3-8 所示。它的原理主要是通过集成物理不可克隆函数到 IC 上来实现。在一个 IC 上喷上一种特殊的涂层，这种涂层包含一些小的随机的介质颗粒。在 IC 顶部的金属层上有电容式传感器，它用来测量介质造成的随机电容。当在 IC 上喷上这种涂层时，不同的电容式传感器就会测量这个涂层造成的随机电容，然后输出涂层 PUF 不同的响应。因此，在涂层 PUF 上，电容式传感器是涂层 PUF 的激励，而介质造成的随机电容是涂层 PUF 的响应。正如光学 PUF 一样，一个涂层 PUF 的激励响应行为也是高度依赖于涂层的微观物理细节的，因此在很大程度上具有唯一性和不可克隆性。其他的模拟电路 PUF 包括类似于涂层 PUF 的 LC－PUF；利用芯片上的模拟测量电路直接测量 MODFET 的阈值电压变化来实现的 PUF，以及基于电阻值来识别一个 IC 电源分配系统方法的 PUF 等。

3. 数字电路 PUF

第三种物理不可克隆函数的实现方法是数字电路 PUF。数字电路 PUF 除了具备物理不可克隆函数本身的条件外，还需要满足两个额外的要求。

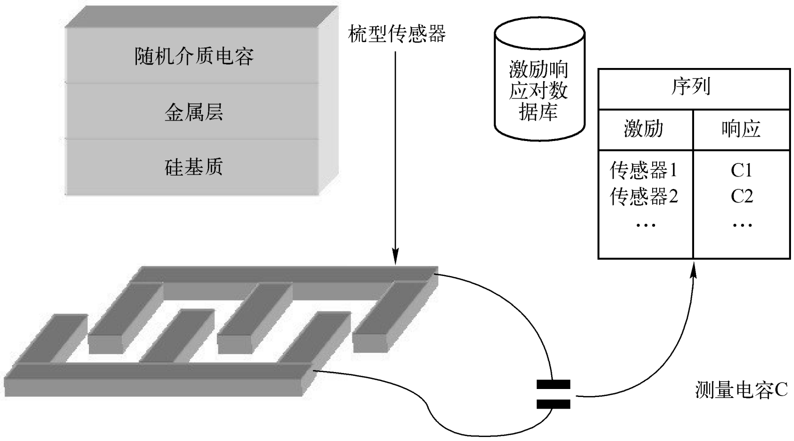


图 3-8　涂层 PUF

首先，一个完整的物理不可克隆函数要完全集成到嵌入式设备中，这里的完整的物理不可克隆函数是指它包括实现物理不可克隆函数的完整设备，如测量设备等。

其次，标准的设备制造流程可以完全地执行这种集成，即不需要特定的物理不可克隆函数过程步骤或组件。

很显然，数字电路 PUF 可能提供更高的安全条件，但也有可能需要更多的实现成本。大部分数字电路 PUF 的实现方法都是基于数字集成电路提出的。其主要的工作原理是，在同样的生产条件下，IC 之间会产生不可避免的随机制造差异，利用这个差异就可以实现数字电路 PUF。目前大部分数字电路 PUF 是在硅器件上实现的，因此也叫硅 PUF。数字电路 PUF 目前主要有两种实现方法。

第一种类型的数字电路 PUF 是利用数字信号的传播延迟变化来实现的。一个数字信号的传播延迟就是信号在路径上遇到元器件电子参数的一个函数。这些元器件的电子参数，如 MOSFET 通道长度、宽度和阈值电压、氧化层厚度、金属线的形状等都会受到制造差异的影响。因此，一个数字信号的传播延迟将会有部分随机性，并且在测量时会显示出类似于物理不可克隆函数的行为。目前有三种主要的具体实现方法，即基于仲裁器的 PUF、基于环形振荡器的 PUF 和毛刺 PUF。基于仲裁器的 PUF 的主要原理是在 IC 上实现两个对称的数字信号延迟路径，电路如图 3-9 所示。一个激励控制着选择路径的确切延迟。引入判决条件是通过两个脉冲同时在两个路径上走，看哪个路径更快并相应由仲裁器电路输出一位响应。因此，它是 n 位激励对应 1 位响应。基于环形振荡器的 PUF 也是通过数字信号延迟差异来实现的，所不同的是它使用负反馈转化数字信号为振荡指标。通过测量振幅，就可以获得测量的延迟。这两种具体的实现方法都是在集成电路上实现和测试的，并且能够显示出一个很好的物理不可克隆函数行为。然而，人们同时认识到由于特定延迟路径的线性构造，两个物理不可克隆函数都易受到模型建立攻击。也就是说，在观察物理不可克隆函数的一些激励响应对以后，敌手可以用很高的概率预测出其余的激励响应对。另一种利用延迟路径差异的物理不可克隆函数称为毛刺 PUF。毛刺的部分随机性是由于不可避免的制造差异引起的，因此通过观察一个随机逻辑电路的毛刺就可以显示出一个物理不可克隆函数的行为。此外，由于毛刺复杂的非线性传播，一个毛刺 PUF 的不可预测性比基于仲裁器的 PUF 和基于环形振荡器的 PUF 更高。

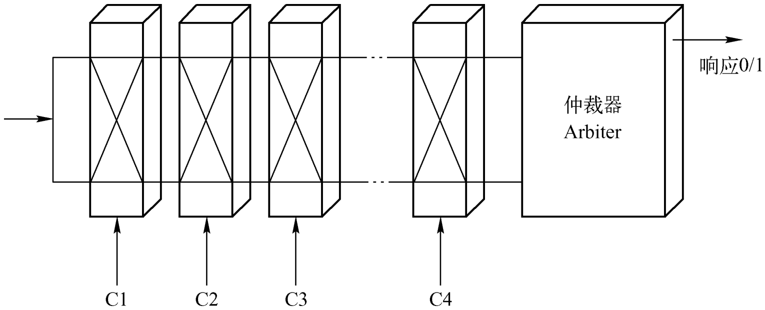


图 3-9　一个基于仲裁器的 PUF 电路

第二种类型的数字电路 PUF 是利用一些存储器单元结构的稳定状态的制造变化来实现的。一般情况下，完成存储器的数字存储是通过双稳态逻辑单元实现的，也就是一个逻辑单元假设有两个不同的但是逻辑上稳定的状态。具体过程是，首先通过交叉耦合两个门器件，如反相器来构建一个双稳态逻辑单元，通过这个双稳态逻辑单元选择寄存在两个中的一个状态，就实现了存储一个二进制数字。但是，如果双稳态逻辑单元进入一个不稳定状态，它就可能在不稳定状态之间振荡，但最后会回到双稳态中的一个，而实验表明大多数单元都会有其明确偏向。这个效果是由其对称设计单元的参数间的不匹配造成的。而这种不匹配是由制造变化差异引起的，因此能够观察到这样一个存储单元的稳定状态显示出一个类似于物理不可克隆函数的行为。因此，一个特定设备的这种稳定的状态是随机和唯一的，并可作为物理不可克隆函数的响应。通过观察一个静态随机存取（Static Random Access Memory，SRAM）单元或一个触发器的稳定状态，就实现了 SRAM PUF（见图 3-10）和触发器 PUF，而锁存 PUF 和蝴蝶 PUF 是通过破坏一个单元之后观察稳定的状态来实现的。总结所有的情况，物理不可克隆函数的激励是一个特定单元的地址，而响应是单元的稳定状态。

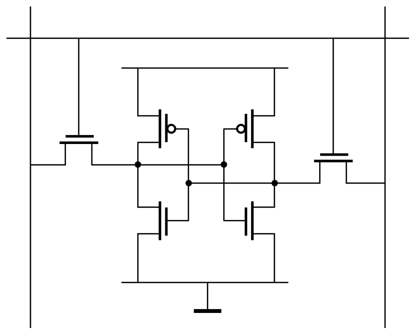


图 3-10　一个 SRAM PUF 单元在 CMOS 电路中

3.3.4　PUF 属性

从物理不可克隆函数实现方法中可以看出，物理不可克隆函数并不是一个单纯的数学概念，而是嵌入了一些物理实体、包含诸多属性及有输入、输出功能的函数。下面给出 PUF 的七个经常出现的属性，并对这些属性进行讨论。在讨论物理不可克隆函数的属性之前，先给出物理不可克隆函数的几个非形式化概念及符号表示。

一般可以将 PUF 看成输入为 C，输出为 R 的函数 P，因此用 $R=P(C)$ 来描述一个特定的物理不可克隆函数的挑战（激励）和响应关系。物理不可克隆函数的基本应用是实现认证。而在基本的认证过程中会遇到错误认证的情况，人们经常借助片间汉明距离与片内汉明距离两个概念来描述这个问题。对于一个物理不可克隆函数，片间汉明距离与片内汉明距离定义如下。

定义 3.3：由于物理不可克隆函数的唯一性和不可克隆性会导致两个不同的物理不可克

隆函数实体产生两个完全不同的响应，因此片间汉明距离是指对两个不同物理不可克隆函数实体输入一个特定的激励后，其产生的两个响应之间的距离。

定义 3.4：在通常情况下，一个物理不可克隆函数响应的精确值不可避免地受到噪声、测量的不确定性和外部因素的影响，因此片内汉明距离是指对一个单一的物理不可克隆函数重复两次输入一个特定的激励后，其产生的响应之间的距离。

从上面的描述可以看出，片间汉明距离与片内汉明距离都是测量激励造成的一对响应之间的距离。在通常情况下，用直方图来表示一个特定类型的物理不可克隆函数的片间汉明距离与片内汉明距离，并且一般用平均值来作为它们的评估指标。μ_{intra}表示平均间距离而μ_{inter}表示平均内距离。因此，通常期望得到的一个物理不可克隆函数的行为是一个尽可能小的μ_{intra}，而μ_{inter}尽可能靠近 50% 。下面给出 PUF 的七个属性。

1. 鲁棒性

物理不可克隆函数具有鲁棒性是指当用激励 x 多次输入同一物理不可克隆函数时，在允许有一小部分错误的条件下，它总是返回相同的响应 $y = P(x)$。这个错误必须在一个很小的考虑距离度量之内。也就是说，这个具有细微差异的响应在距离度量中是非常靠近的。物理不可克隆函数的鲁棒性主要通过片内汉明距离直方图来量度并由其实现结果的平均值μ_{intra}来体现。鲁棒性是物理不可克隆函数和伪随机数发生器（Pseudo－Random Number Generator，PRNG）的本质区别属性。在所有已提出的 PUF 实现方法中，几乎所有的物理不可克隆函数都具有鲁棒性。

2. 可计算性

物理不可克隆函数具有可计算性是指给定物理不可克隆函数 PUF 和激励 x，可以很容易地计算出相应的响应 $y = P(x)$。这里可以从不同的角度来理解。从理论的角度来看，它意味着在多项式时间和资源内计算是可行的。而从实际的角度来看，它意味着一个非常低的成本开销，即在有限的时间、空间、功耗和集成芯片的能源等约束条件下，计算是可行的。此外，如果一个物理不可克隆函数是可计算的，则表明运行一个物理不可克隆函数是可行的。从这个意义上说，所有能提供实验结果的 PUF 实现方法至少在理论上是可计算的。在所有已提出的 PUF 实现方法中，光学 PUF 和涂层 PUF 等一些非数字电路 PUF 特殊的构造步骤需要额外的制造步骤或外部测量设备才能满足可计算性。SRAM PUF 需要设备电源才能满足可计算性。而其他大多数的 PUF 实现方法都是满足可计算性的。可以说可计算性仅仅是一个实用性的约束。

3. 唯一性

物理不可克隆函数具有唯一性是指物理不可克隆函数的响应包含了物理实体嵌入物理不可克隆函数 PUF 的一些身份信息。从信息理论的角度来看，唯一性意味着在物理不可克隆函数全体中，一个特定的物理不可克隆函数的一些 CRP 满足于唯一标识。也就是说，在理想情况下，一个物理不可克隆函数的响应会对全体造成多个分区，这样一来，连续的响应就会使全体的分区越来越小，直到优化了整个分区。这样，CRP 集合就可以唯一认证一个物理不可克隆函数的实现。在大多数的实验中，主要通过片间汉明距离直方图来测量唯一性，

并且通过其平均值 μ_{intra} 来体现。在已提出的 PUF 实现方法中，所有的物理不可克隆函数都具有唯一性。

4. 不可克隆性

不可克隆性是物理不可克隆函数的根本属性，这从“不可克隆”函数的命名上就可以看出。不可克隆性其实是一个程序过程，这个过程既可以是物理过程，也可以是数学过程，因此不可克隆性可以分为物理不可克隆性和数学不可克隆性。一个物理不可克隆函数具有物理不可克隆性是指，给定物理不可克隆函数 PUF，构造一个物理实体包含另一个物理不可克隆函数 P，使得对于任意 $x \in X$：在很小错误的情况下 $P'(x) = P(x)$ 是困难的。从定义可以看出，物理不可克隆性表示一个敌手物理克隆一个物理不可克隆函数的困难性。因为产生一个物理克隆的难度甚至可以达到原始物理不可克隆函数 PUF 的制造者也很难实现的程度，所以它也被称为制造者阻力。而一个物理不可克隆函数具有数学不可克隆性是指，给定一个物理不可克隆函数 PUF，构造一个数学程序 f' 使得对于任意 $x \in X$：在很小错误的情况下 $f'(x) = P(x)$ 是困难的。因为有些实现方法可以容易实现物理克隆而不是数学克隆或反之亦然，所以物理和数学不可克隆性是两个根本不同的属性。因此，为了真正实现不可克隆性，需要同时实现物理和数学不可克隆性。在所有的 PUF 实现方法中，几乎所有的物理不可克隆函数都具有物理不可克隆性，但数学不可克隆性却很少满足。对于基于仲裁器的 PUF 和一部分基于环形振荡器的 PUF 来说，通过使用模型建立攻击很容易破解数学不可克隆性。对于涂层 PUF 和 SRAM PUF 来说，通过搜集很多 CRP，也可以很容易破解数学不可克隆性。

5. 不可预测性

物理不可克隆函数具有不可预测性是指给定一个挑战响应对集合 $\text{CRP} = \{x_k, y_k = P(x_k)\}$，很难在一个很小的错误范围内预测响应 $P(x_i)$，其中 x_i 是一个随机激励且 $(x_i, y_i) \notin \text{CRP}$。从这个定义中可以看出，不可预测性是不可克隆性的松散形式，即不可克隆性意味着不可预测性。在所有的 PUF 实现方法中，大多数物理不可克隆函数的实现方法都具有不可预测性。对于基于仲裁器的 PUF 和一部分基于环形振荡器的 PUF 来说，通过学习算法可以预测出物理不可克隆函数的新的 CRP。也就是说，敌手在学习一个物理不可克隆函数大量 CRP 的基础之上，有可能预测出一个新的 CRP，从而消除物理不可克隆函数的不可预测性。

6. 轻量级

物理不可克隆函数具有轻量级属性是指实现物理不可克隆函数 PUF 元器件的数量和大小都是很小的。这在资源有限的设备中有广大的应用前景。例如，在智能卡、RFID 和传感器网络节点等资源有限的设备中，一些成熟的加密算法因为器件太多或消耗过大等原因不能使用，所以轻量级就成为硬性的要求。到目前为止，已经提出多种基于物理不可克隆函数轻量级的认证计划和密钥生成的应用。在所有的 PUF 实现方法中，数字电路 PUF 更具有这个属性，因为它是利用数字电路内在的变化，而不需要特殊编程实现具体的电路。而且在密钥生成的一些应用中，物理不可克隆函数的响应不需要存储而是直接传送到电路的输入中进行下面的计算，这也正是运用了物理不可克隆函数的这个属性。

7. 防篡改

物理不可克隆函数具有防篡改属性是指当把改变的物理实体嵌入物理不可克隆函数 PUF 使得 $P \rightarrow P'$ 时，有非常高的概率 $\exists x \in X : P(x) \neq P'(x)$。从这个定义可以看出，防篡改属性是指在篡改发生之后检测篡改攻击的能力。由于物理不可克隆函数依赖于微小的物理构造差异，所以人们通常认为篡改一个物理不可克隆函数将不可避免地改变物理不可克隆函数的激励响应行为。这意味着，对物理不可克隆函数的篡改攻击将对 CRP 行为造成不可消除的痕迹。在所有的 PUF 实现方法中，只有光学 PUF 和涂层 PUF 是明确防篡改的，其他的物理不可克隆函数构造是否具有防篡改属性，还有待研究。

3.3.5　PUF 研究及应用现状

PUF 技术从 2001 年提出以后，逐渐引起了学术界的注意，近年来已经有一些关于 PUF 设计实现、PUF 芯片及电路防伪、PUF 认证技术的研究，也出现了很多这方面的论文和资料。但总体来说 PUF 还是非常新颖的内容，相关研究并不是很多，尤其是在国内，PUF 技术的研究基本上属于空白阶段，很少有文献和资料对其进行介绍。

目前国际上对 PUF 技术研究处于领先地位的是美国的 Verayo（威诚伟特）公司，目前该公司已经有部分技术应用于 RFID 射频识别系统，有效地实现了对标签的安全性认证，很好地解决了射频标签伪造问题。随着时间的推移，人们已经基于各种实现方法提出了物理不可克隆函数更加多样性的应用。这些应用归纳起来可基本上分为 IP 保护、系统认证和密钥生成等。在这些应用环境中，物理不可克隆函数主要有三种用法：认证、密钥生成器和可计算函数。

1. 认证

认证是物理不可克隆函数最基本的应用。由于物理不可克隆函数的不可克隆性、防篡改和轻量级等属性，使用物理不可克隆函数用于认证是一种非常有用的防伪技术。因此，在物理不可克隆函数的相关应用文献中，这是最常见的形式。它的基本原理是这样的：在注册阶段，每一个物理不可克隆函数的一些 CRP 连同嵌入物理不可克隆函数的物理系统的身份一起被存储在数据库中，在认证阶段，验证者从数据库中挑选一个随机 CRP，然后提供给当前的系统来激励物理不可克隆函数，如果观察到物理不可克隆函数的响应足够接近于数据库中存储的响应，则认证成功，否则失败。为了防止重放攻击，每个物理不可克隆函数的每个 CRP 只能使用一次并且必须在验证结束后从数据库中删除。

在这样一个过程中会遇到错误认证的问题。认证错误主要包括两种类型：第一种是错误接受，即通过认证接受了错误的物理不可克隆函数；第二种是错误拒绝，即通过认证拒绝了正确的物理不可克隆函数。一个认证正确与否的阈值分别取决于片内汉明距离直方图和片间汉明距离直方图。如果两个直方图不重叠，就可以通过在两个直方图之间差距的某处选择阈值来进行无差错的认证，如图 3-11（a）所示。但是如果由于设备老化等原因使得它们重叠，则设置的阈值就必须在错误接受和错误拒绝之间权衡，如图 3-11（b）所示。从图 3-11 中可以看出，最佳的选择是通过在两个分布图交集中的某处设置阈值来最大限度地降低错误接受和错误拒绝的总和，但具体的权衡取决于具体的应用。

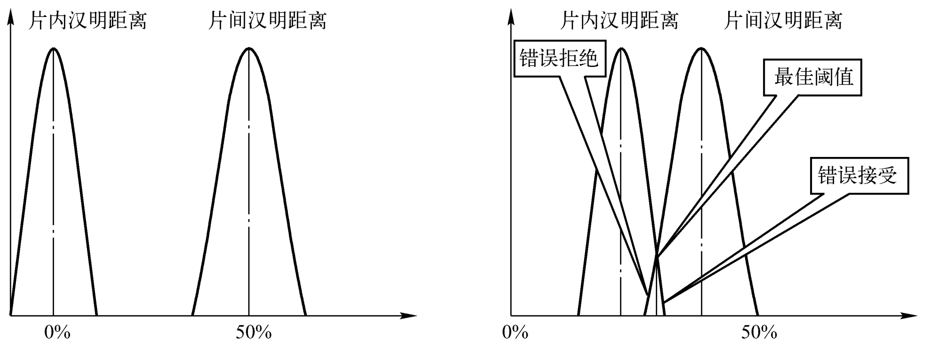


图 3-11　物理不可克隆函数认证中片内汉明距离和片间汉明距离的关系

2. 密钥生成器

基于物理不可克隆函数的密钥生成在安全领域的应用已经有大量的文献。应用物理不可克隆函数作为密钥生成器主要考虑使用的是数字电路 PUF。因为在集成电路中，数字电路 PUF 有很好的属性用于密钥生成和存储。通过采用适当的后处理技术，物理不可克隆函数可用于生产一个加密强密钥，使得方案更具有安全性。由于物理不可克隆函数的一些有用属性，使用物理不可克隆函数生成密钥有以下五个方面的优势。

① 物理不可克隆函数可能的防篡改属性可以用来提供防篡改的密钥存储。

② 由于随机性是永久固定在细微芯片的物理构造上的，所以不需要传统的非易失性存储步骤。这也额外地提供了对探测攻击和其他可能的侧信道攻击的安全性。物理不可克隆函数可以在规定的时间内派生出安全密钥并且在使用之后删除，因此密钥不需要永久保存成数字格式，而只是当需要操作时出现在非易失性存储器中。这就限制了提取设备中密钥的攻击时限。

③ 一个用物理不可克隆函数生成的密钥也是密切相关于嵌入物理不可克隆函数的物理硬件，使得整个硬件具有物理不可克隆性。

④ 在安全存储一个加密密钥方面，使用具有模糊提取模块的物理不可克隆函数比使用非易失性存储单元更有效。

⑤ 由于密钥生成的数字电路随机性是由不可避免的制造变化引起的，所以不需要明确的密钥编程步骤，这简化了密钥分配。

从物理不可克隆函数响应中提取一个安全的密钥需要处理两个主要问题：首先，在不同的测量中，所有数字电路 PUF 实现方法产生的响应都有一个非负概率的错误，因此，在后处理过程中就需要采用一个纠错步骤来保证每次派生出相同的密钥；其次，提取算法需要确保输出的密钥是完全不可预测的，也就是说，它应该是一个均匀分布的随机比特串。由于物理不可克隆函数的响应大多数只有部分是不可预测的，所以提取算法需要压缩一些响应到一个密钥中，以保证强的不可预测性。

文献“Fuzzy extractors：How to generate strong keys from biometrics and other noisy data”已经给出同时满足这两个要求的算法，称为模糊提取算法，如图 3-12 所示。模糊提取的主要想法是：在最初的生成阶段，给物理不可克隆函数输入一个激励并产生一个响应，

然后模糊提取算法根据响应产生一个包含额外信息的密钥。这些额外信息通常被称为辅助数据。这两个数据都被验证者存储在一个安全的数据库中而不在设备上。在认证阶段，验证者把辅助数据提供给算法，算法用它来从物理不可克隆函数中提取相同的密钥，具体过程如图 3-13 所示。这样一来，含有物理不可克隆函数的设备和验证者之间就建立了一个共享密钥。



图 3-12　模糊提取

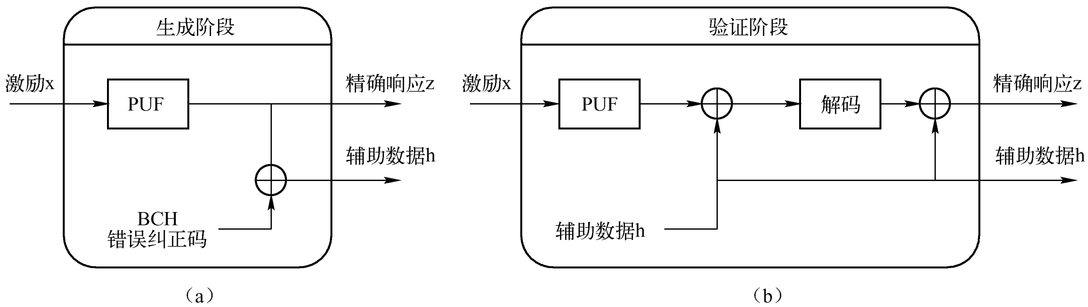


图 3-13　传统的模糊提取算法

3. 可计算函数

由于基于延迟的数字电路 PUF 可以采用线性不等式来表示，所以它可以用作一个可计算函数。这意味着服务器不需要存储 CRP，而是直接计算出预计的响应。这就使得物理不可克隆函数具有轻量级的优势，更适合在像智能卡、RFID 和传感器网络节点等资源有限的设备中使用。但是线性结构带来优势的同时也带来了威胁，即敌手很容易通过这个线性结构数学克隆这个物理不可克隆函数。因此，这类物理不可克隆函数实现方法的安全性主要面临两个挑战。

第一个挑战是抵抗建模攻击的问题。在一个标准的攻击模型中，在通过学习这个物理不可克隆函数的一些激励响应对之后，给定敌手一个新的激励，则他应该无法预测出相应的响应。文献“A Tamper－Proof and Lightweight Authentication Scheme”通过使用线性规划等数学方法，给出了模拟基于仲裁器的 PUF 的一个线性模型。这个线性模型本质是一种激励响应数据库的压缩版本。通过这个线性模型，模拟这个物理不可克隆函数理论上可以达到任何级别的精度。虽然理论上模拟这类物理不可克隆函数可以达到任何级别的精度，但是在实际的应用中温度、电压和电磁泄漏等外部噪声对模拟效果会产生较大影响。

第二个挑战是轻量级的问题，即实现安全性的同时要尽可能地使用低开销，因此在资源有限和算法安全性上要做出权衡。在智能卡、RFID 和传感器网络节点等资源有限的设备上，使用物理不可克隆函数作为一个可计算函数的一个主要出发点是轻量级属性。因为在这些资源有限的设备上，传统的加密算法和认证协议都消耗太大，不能满足能量小和覆盖范围大的要求。而且哈希函数和物理不可克隆函数相比也是高消耗的。除了其轻量级属性外，物理不可克隆函数还具有防篡改、不可克隆性等一些好的属性，因此用物理不可克隆函数来加强资

源有限设备安全性是一个有效的方法。

参考文献

[1] Gassend B, Clarke D, Dijk M, et al. Controlled Physical Random Functions [C]. In: Proceedings of 18th Annual Computer Security Applications Conference, December 2002: 149 – 160.

[2] Gassend B, Clarke D, Dijk M, et al. Silicon Physical Random Functions [C], In: Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS02), November 2002: 148 – 160.

[3] Lee J W, Lim D, Gassend B, et al. A Technique to Build A Secret Key in Integrated Circuits with Identification and Authentication Applications [C]. In: Proceedings of the IEEE VLSI Circuits Symposium, June 2004: 176 – 179.

[4] Skoric B, Schrijen G – J, Ophey W, et al. Experimental Hardware for Coating PUFs and Optical PUFs [J] Security with Noisy Data, Springer London, 2007: 255 – 268.

[5] Maes R, Tuyls P, Verbauwhede I, Low – Overhead Implementation of a Soft Decision Helper Data Algorithm for SRAM PUFs [C], In: Proc. CHES2009, 2009: 332 – 347.

[6] Suh G E, Devadas S, Physical Unclonable Functions for Device Authentication and Secret Key Generation [C], in Proc. DAC, 2007: 9 – 14.

[7] KumarS, Guajardo J, Maes R, et al. The Buttefly PUF: Protecting IP on Every FPGA [C], In IEEE International Workshop on Hardware Oriented Security and Trust, Anaheim 2008: 67 – 70.

[8] Guajardo J, Kumar S S, Schrijen G – J, et al. FPGA Intrinsic PUFs and Their Use for IP Protection [C]. In P. Paillier and I. Verbauwhede, editors, Cryptographic Hardware and Embedded Systems (CHES 2007), 47 (27), September 2007: 63 – 80.

[9] Guajardo J, Kumar S S, Schrijen G – J, et al. Physical Unclonable Functions and Public KeyCrypto for FPGA IP Protection [C]. In: International Conference on Field Programmable Logic and Applications (FPL 2007), August 27 – 30, 2007: 189 – 195.

[10] Holcomb D E, Burleson W P, Fu K, Initial SRAM State As A Fingerprint And Source of True Random Numbers For RFID Tags. Conference on RFID Security 07, July 11 – 13, 2007.

[11] Cortese Pier Francesco, Gemmiti Francesco, Palazzi Bernardo, et al. Efficient and Practical Authentication of PUF – Based RFID Tags in Supply Chains [C], Program for the IEEE International Conference on RFID – Technology and Applications, Guangzhou, China, June 2010: 182 – 188.

[12] Kulseng L, Yu Z, Wei Y, et al. Lightweight Mutual Authentication and Ownership Transfer for RFID Systems [C], In: Proc. INFOCOM, 2010: 251 – 255.

[13] Tuyls P, Batina L, RFID – tags for Anti – Counterfeiting, Topics in Cryptology CT – RSA [C], Lecture Notes in Computer Science, Vol. 3860, San Jose, CA, 2006: 115 – 131.

[14] Oliveira L B, Scott M, Lopez J, et al. TinyPBC: Pairings For Authenticated Identity – Based Non – Interactive Key Distribution In Sensor Networks [C], In Proc. of International Conference Networked Sensing Systems INSS2008, 2008: 173 – 180.

[15] PAPPU RS. Physical one – way functions [D]. Boston: Massachusetts Institute of Technology, 2001.

[16] TUYLS P, SCHRIJEN GJ, KORIC B, et al. Read – proof hardware from protective coatings [A]. In: Cryptographic Hardware and Embedded Systems workshop [C]. Yokohama Japan: springer, 2006: 369 – 383.

[17] LIM D. Extracting Secret Keys from Integrated Circuits [D]. Boston: Massachusetts Institute of Technology, 2004.

[18] LEE JW, LIM D, GASSEND B, et al. A Technique to Build a Secret Key in Integrated Circuits for Identifica-

tion and Authentication Application [A]. In: Proceedings of the Symposium on VLSI Circuits [C]. Honolulu USA: Digest of Technical Papers, 2004: 176 – 159.

[19] KUMAR S, GUAJARDO J, MAES R, et al. The Butterfly PUF Protecting IP on Every FPGA [A]. In: IEEE International Workshop on Hardware – Oriented Security and Trust [C]. Anaheim CA USA: IEEE, 2008: 67 – 70.

[20] GUAJARDO J, KUMAR SS, SCHRIJEN GJ, et al. FPGA Intrinsic PUFs and Their Use for IP Protection [A]. In: Cryptographic Hardware and Embedded Systems Workshop [C]. Vienna Austria: Springer, 2007: 63 – 80.

[21] GASSEND B, CLARKE D, VANDIJK M, et al. Delay – based Circuit Authentication and Applications [A]. In: ACM Symposium on Applied Computing [C]. New York: ACM Press, 2003: 294 – 301.

[22] HAMMOURI G, OZTURK E, SUNAR B. A Tamper – Proof and Lightweight Authentication Scheme [J]. Pervasive and Mobile Computing. 2008. 4 (6): 807 – 818.

[23] OZTURK E, HAMMOURI G, SUNAR B. Towards robust low cost authentication for pervasive devices [A]. In: Proceedings of the Sixth IEEE International Conference on Pervasive Computing and Communication [C]. Hong Kong: IEEE Computer Society, 2008: 170 – 178.

[24] SADEGHI AR, SCHULZ S, WACHSMANN C. Lightweight Remote Attestation using Physical Function [A]. In: Conference on Wireless Network Security [C]. Hamburg Germany: ACM press, 2011: 109 – 114.

[25] TUYLS P, BATINA L. RFID – tags for anti – counterfeiting [J]. RSA 2006 conference. San Jose USA, 2006: 13 – 17.

[26] BATINA L. GUAJARDO J. KERINS T, et al. Public – key cryptography for rfid – tags [A]. In: Pervasive Computing and Communications Workshops [C]. white plains USA: IEEE Computer Society, 2007: 217 – 222.

[27] GASSEND B, CLARKE D, VANDIJK M, et al. Silicon physical random functions [A]. In: ACM Conference on Computer and Communications Security [C]. New York USA: ACM, 2002: 148 – 160.

[28] GASSEND B. Physical Random Functions [D]. Massachusetts: Massachusetts Institute of Technology, 2003.

[29] G. E Suh and S Devadas. Physical unclonable functions for device authentication and secret key generation [A]. In: Design Automation Conference [C], 2007: 9 – 14.

[30] C – E. D Yin and G Qu. Maximizing RO PUF' s secret extraction [A]. In: IEEE Symposium on Hardware – Oriented Security and Trust [C], 2010: 100 – 105.

[31] A Maiti, J Casarona, L McHale, and P Sxhaumont. A Large scale characterization of RO – PUF [A]. In: IEEE Symposium on Hardware – Oriented Security and Trust [C]. 2010: 94 – 99.

[32] E Ozturk, G Hammouri, and B Sunar. Physical Unclonable Function with tristate buffers [A]. In: IEEE Symposium on Circuits and Systems [C], 2008: 3194 – 3197.

[33] L Lin, D Holcomb, D. K Krishnappa, P Shabadi, and W Bureson. Low – power sub – threshold desigh of secure physical unclonable functions [A]. In: ACM IEEE international symposium on Low power electronics and design [C]. 2010: 43 – 48.

[34] MAJZOOBI M, KOUSHANFAR F, POTKONJAK M. Techniques for design and implementation of secure reconfigurable pufs [J]. ACM Transaction on Reconfigurable Technology System, 2009, 2 (1): 1 – 33.

[35] U Ruhrmair, F Sehnke, J Solter, G. Dror, S Devadas, and J Schmidhuber. Modeling attacks on Physical Unclonable Functions [A]. In: ACM conference on Computer and Communications security [C], 2010: 237 – 249.

[36] HOLCOMB DE, BURLESON WP, FU K. Initial SRAM State as a Fingerprint and Source of True Random Numbers for RFID Tags [A]. In: Proceedings of the Conference on RFID Security [C]. Malaga Spain:

RFID Publications, 2007: 11 –13.

[37] MAES R, TUYLS P, VERBAUWHEDE T. Intrinsic PUFs from Flip – Flops on Reconfigurable Devices [A]. In: 3rd Benelux Workshop on Information and System Security [C]. Eindhoven the Netherlands, 2008: 17.

[38] V van der Leest, G – J Schrijen, H Handschuh, and P Tuyls. Hardware intrinsic security from D flip – flops [A]. In: ACM Workshop on Scalable Trusted Computing [C], 2010: 53 –62.

[39] SU Y, HOLLEMAN J, OTIS B A. 1.6pj/bit 96% Stable Chip – ID Generating Circuit Using Process Variations [A]. In IEEE International Solid – State Circuits Conference [C]. Washington DC, IEEE Computer Society, 2007: 406 –611.

[40] BULENS P, STANDAERT FX, QUISQUATER JJ. How to Strongly Link Data and Its Medium, The Paper Case [J]. IET Information Security. 2010, 4 (3): 125 –136.

[41] J. DR Buchanan, R. P. Cowburn, A. – V. Jausovec, D Petit, P Seem, G Xiong, D Atkinson, K Fenton, D. A Allwood, and M. T Bryan. Forgery: fingerprinting documents and packaging [J]. Nature, Vol. 436, No. 7050: 475, 2005.

[42] HAMMOURI G, DANA A, SUNAR B. CDs Have Fingerprints Too [A]. In Proceedings of the 11th International Workshop on Cryptographic Hardware and Embedded Systems [C]. Berlin, Heidelberg: Springer, 2009: 348 –362.

[43] GUAJARDO J, KORIC B, TUYLS P, et al. Anti – counterfeiting, key distribution, and key storage in an ambient world via physical unclonable function [J]. Information Systems Frontiers, 2009, 11 (1): 19 –41.

[44] LOFSTROM K, DAASCH WR, TAYLOR D. IC Identification Circuit Using Device Mismatch [A]. In: Proceedings of Solid – State Circuits Conference [C]. San Francisco CA: IEEE, 2000: 372 –373.

[45] HELINSKI R, ACHARYYA D, PLUSQUELLIC J. A physical unclonable function defined using power distribution system equivalent resistance variation [A]. In Proceedings of the 46th Annual Design Automation Conference [C]. New York USA: ACM, 2009: 676 –681.

[46] PELGROM M, DUINMAIJER A, WELBERS A. Matching properties of mos transistors [J]. IEEE Journal of Solid – State Circuits, 1989, 24 (5): 1433 –1439.

[47] AGARWAL A, KANG K, BHUNIA S, et al. Device – aware yield – centric dual – vt design under parameter variations in nanoscale technologies [J]. IEEE Transactions on Very Large Scale Integration System, 2007, 15 (6): 660 –671.

[48] KANAMOTO T, OGASAHARA Y, NATSUME K, et al. Impact of well edge proximity effect on timing [J]. IEEE Journal of Solid – State Circuits, 2008, 91 – A (12): 3461 –3464.

[49] REDA S, NASSIF SR. Analyzing the impact of process variations on parametric measurements: novel models and applications [A]. In: Proceedings of the Conference on Design, Automation and Test [C]. Nice France: IEEE, 2009: 375 –380.

[50] WESTE NHE, HARRIS DM. CMOS VLSI Design: A Circuits and Systems Perspective [D]. Boston, Massachusetts: Pearson Education, 2010.

[51] SUZUKI D, SHIMIZU K. The glitch puf: A new delay – puf architecture exploiting glitch shapes [A]. In: Crypt – ographic Hardware and Embedded Systems [C]. Berlin eidelberg: Springer, 2010: 366 –382.

第 4 章　感知层认证技术

物联网感知层认证技术是物联网安全技术的重要组成部分之一。通过认证技术可以证实被认证对象的身份是否属实、消息是否可信。相对于互联网的认证技术，物联网感知层认证技术需要具有更低的计算、通信及存储开销，以满足终端设备资源受限的要求。本章在对物联网感知层认证技术介绍的基础上，针对物联网感知层安全认证的需求及网络特点，重点分析讨论了 RFID 及无线传感器网络的相关认证方法及协议，给出了相关的典型认证机制及常见协议。

4.1　感知层认证技术概述

由于物联网终端设备往往分布广泛，常常部署在野外、战场等管理人员难以到达的区域，因此很难进行安全维护。自身资源的受限加之应用环境的恶劣，导致物联网终端设备面临严重的安全威胁，如由于管理人员难以对部署在野外环境的传感器节点进行监测和维护，因此攻击者便很容易实现对传感器节点的捕获，并通过分析器内部数据伪造更多的“合法”节点，利用身份的欺骗来实现对目标网络的攻击或散步虚假的消息。针对物联网感知层中典型的 RFID 网络和无线传感器网络而言，这种类似于身份欺骗和信息欺骗的攻击有很多，表 4-1、表 4-2 分别列出了 RFID 和无线传感器网络中的部分欺骗攻击方法，这些攻击给网络安全带来了极大的威胁。

表 4-1　RFID 安全威胁

名　　称	解　　释
伪造攻击	指伪造电子标签产生系统认可的“合法用户标签”
假冒攻击	指在射频通信网络中，攻击者截获一个合法用户的身份信息后，利用这个身份信息来假冒该合法用户接入网络
复制攻击	通过复制他人的电子标签信息，多次顶替他人使用
重放攻击	指攻击者用某种方法将用户的某次使用过程或身份验证记录重放或将窃听到的有效信息经过一段时间以后再传给信息接收者，骗取系统的信任
信息篡改	指攻击者将窃听到的信息进行修改之后再将信息传送给接收者

表 4-2　无线传感器网络安全威胁

名　　称	解　　释
伪造攻击	指攻击者伪造传感器节点，产生网络认可的“合法用户”
假冒攻击	指在传感器网络中，攻击者截获一个合法用户的身份信息后，利用这个身份信息来假冒该合法用户接入网络。Sybil 攻击就是攻击者通过假冒多个合法节点的身份，骗取其他节点的信息，使之成为网络的一员，并配合其他攻击手段进行的一种假冒攻击
虚假路由	通过欺骗、篡改或重发路由信息，攻击者可以造成路由环路，或抵制信息传输，延长或缩短路径，实现网络分割、资源消耗等
重放攻击	攻击者使节点误认为加入了一个新的会议，并截获网络中传播的传感信息、控制信息、路由信息等，再对这些截获的旧信息进行重新发送，造成网络混乱

为了抵抗这种欺骗性的网络攻击，一个比较有效的安全防护机制就是认证技术。通过认证，可以实现对用户身份合法性的确认，可以实现对消息来源及完整性的确认，同样也可以实现对目标行为合法性的确认。认证技术是物联网安全的重要技术手段，是防止网络欺骗攻击的主要方法，一般分为两种：

- 消息认证，保证消息的完整性和抗否认性，防止消息在传输过程中被非法篡改，防止消息源对发送消息的抵赖；
- 身份认证，鉴别用户身份，识别和区分访问者的身份，或验证被访问者所声明的身份。

4.1.1　RFID 认证技术

RFID 认证协议主要用于实现阅读器和射频标签之间的安全身份及消息认证。传统计算机网络、互联网中的认证技术非常成熟，但这些认证技术往往基于强大的终端设备和良好稳定的通信链路，对于资源受限的 RFID 系统来说很难直接应用。到目前为止，国内外学者对 RFID 系统的认证技术进行了大量的研究，取得了一定的研究成果，其中比较有代表性的主要有基于 Hash 函数的 Hash – Lock 认证协议、随机 Hash – Lock 认证协议、Hash 链协议和分布式询问 – 应答认证协议，当然也有基于硬件特征的，如基于物理不可克隆函数（PUF）的认证协议等。根据协议所需要计算资源的多少，通常可以将 RFID 认证协议分为三类：重量级、中量级和轻量级协议，如图 4-1 所示。后文中将会对适用于物联网 RFID 系统的比较典型的中轻量及典型协议进行详细介绍。

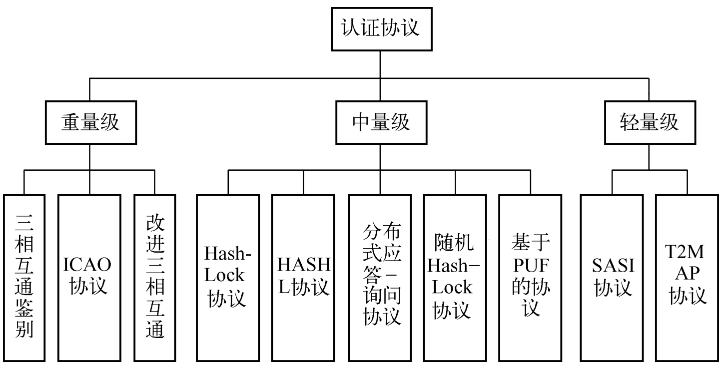


图 4-1　RFID 系统安全认证协议分类

4.1.2　无线传感器网络认证技术

物联网感知层无线传感器网络从体系结构上来说相对 RFID 系统要复杂一些，大量传感器节点具有自组成网的能力，节点间的数据传输也有着点对点、广播、多播等多种通信方式，其安全也面临着更复杂的情况，因此无线传感器网络不但需要对节点和用户的身份进行认证，也需要对点对点、广播、组播等通信信息进行消息认证。从目前的研究情况来看，传统的静态无线传感器网络的认证技术研究相对成熟，而动态无线传感器网络和延迟容忍无线传感器网络的认证研究还有待进一步深入。

在身份认证方面，按照采用的加密形式可以分为四种，如图 4-2 所示。由于传感器节点能量有限的特点，致使很多计算量、通信量大的认证算法无法在无线传感网上使用，所以

对称密码算法的认证机制具有计算、通信和存储开销小、加密速度快、加密效率高的特点，非常适合传感器网络。公钥认证机制往往需要加大通信、计算和存储开销，较难应用于传感器网络，不过由于公钥机制在密钥管理方面的优势仍然吸引着大批学者进行优化设计，尤其是针对椭圆曲线公钥算法的优化研究有着非常好的应用前景；随着技术的进步，相信公钥算法也可以在实际传感器网络认证机制中使用。基于“秘密共享”的认证机制并不使用加密手段，而是用“秘密”的概念作为认证的基础，比较典型的秘密共享认证机制是 K. Bauer 等提出的一种分布式认证协议，这种协议虽然没有加解密的计算开销，但却有着通信开销大、网络延迟长的缺点。动态用户认证同样不采用加解密算法，但却使用逻辑异或操作和单向散列函数等操作，Wong 等人提出了动态的无线传感器网络用户认证机制，该机制的计算开销同样非常大，且步骤烦琐，难以适应于传感器网络。

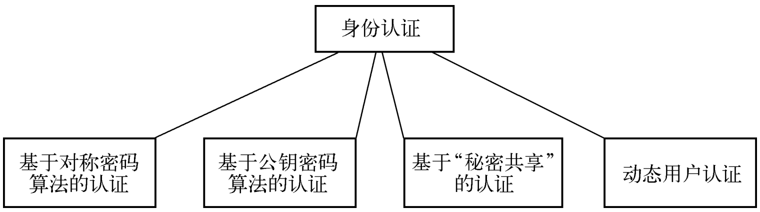


图 4-2　传感器网络身份认证机制分类

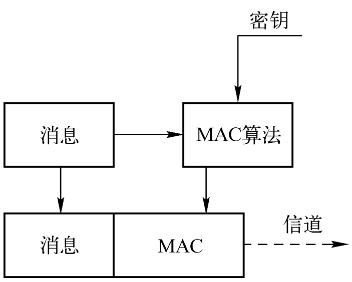


图 4-3　MAC 生成示意图

消息认证在无线传感器网络中应用得最为频繁，无线传感器网络通常使用对称密码的消息认证码（MAC）机制实现消息认证，MAC 生成示意图如图 4-3 所示。消息接收方通过对 MAC 的验证可以确定消息的可信性。MAC 可以非常有效地应用于点到点的双向认证中，但却不能直接用于广播认证，因为 MAC 算法是对称的，发送方和接收方共享秘密密钥，任何一个接收方都能假装成发送方发送伪造消息。目前无线传感器网络中消息认证机制的主要代表有 SNEP、Tinysec、sensec、Minisec 和 uTESLA 等。前四种机制都针对网络中的单播消息，通过 MAC 计算实现了消息的机密性、可认证性、完整性和新鲜性。uTESLA 针对网络中广播消息的认证问题，通过延迟对称密钥的发布，从而引入非对称性，实现有效的广播认证。同样，PUF 作为一种较为新颖的安全单元，不但可以应用于 RFID 系统中，也可以应用于无线传感器网络中，后面将对典型的 SNEP、uTESLA 及基于身份加密及 PUF 的认证机制进行介绍。

4.2　RFID 认证机制

4.2.1　基于 Hash 函数的认证机制

1. Hash – Lock 认证协议

为避免信息泄露和被追踪，2003 年 Sarma 等人提出了 Hash – Lock 认证机制，主要是通

过哈希函数去设定标签锁定或解除锁定。该机制中，射频标签只对授权的阅读器起作用，阅读器对每一个电子标签都有一个认证密匙 k，每个电子标签都存储有一个 Hash 函数计算的结果 metaID = Hash(k)，metaID 用来代替真实的标签 ID；后台数据库内存放着每个标签的认证密钥 k，并且会对应到标签所存储的 metaID，以 metaID 与 k 值作为判断标签锁定及解锁的依据。其中标签状态为锁定时表示阅读器只能够读取到标签的一部分资料，合法的阅读器可以通过这部分资料到后台数据库找出解锁的密钥 k，当标签验证解锁密钥 k 正确后，标签的状态会由锁定转换为解除锁定的状态，此时阅读器就可以读取标签上的所有资料了。其协议流程如图 4-4 所示。

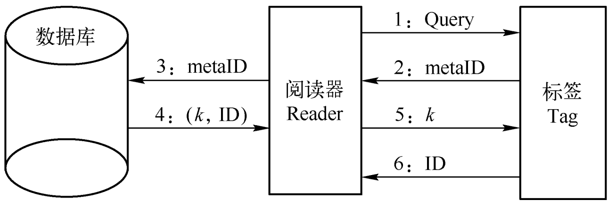


图 4-4　Hash - Lock 认证流程

Hash - Lock 协议的具体执行过程如下。

① 阅读器向标签发送 Query 认证请求。

② 标签将自身内部存储的 metaID 发送给阅读器。

③ 阅读器将 metaID 转发给后台数据库。

④ 数据库查询自身存储的标签数据，如果找到与 metaID 对应的项，则将该项对应的（k,ID）发送给阅读器，其中 ID 为待识别标签的唯一标识符；若无法找到对应的项，则通知阅读器，认证失败。

⑤ 阅读器将接收的来自数据库的信息 k 发送给标签。

⑥ 标签验证 metaID = Hash(k)是否成立，若成立则将其 ID 发送给阅读器。

⑦ 阅读器将从数据库接收到的 ID 与从标签接收到的 ID 相比较，如果一致，则认证通过，否则认证失败。

从上述过程可以看出，Hash - Lock 协议是一种低成本安全与隐私问题解决方法，但该协议中没有 ID 动态刷新机制，并且 metaID 也保持不变，ID 以明文的形式通过不安全的信道传送，因此它很容易受到假冒攻击和重传攻击，攻击者也可以很容易地对标签进行追踪。总体而言，Hash - Lock 协议没有达到其安全目标，无法实现 RFID 系统的安全性和认证需求。

2. 随机化 Hash - Lock 认证机制

为了避免被跟踪，射频标签的响应应是不能被预测到的，而且应该是随机的。主要有两种随机化 Hash - Lock 认证协议。一种是 2003 年 Weis 等提出的随机化 Hash - Lock 认证协议，另一种是 Hash 链协议。

随机化 Hash - Lock 认证协议采用了基于随机数的询问 - 应答机制。当阅读器向射频标签发出 ID 访问请求时，标签向阅读器发出的不是固定的 metaID，而是变化的应答数据。每个标签与阅读器共享一个认证密钥 ID_k。当阅读器向射频标签发出 ID 访问请求时，射频标签

产生一个伪随机数字 R 和输出$(R, H(ID_k \parallel R))$，其中 $H(ID_k \parallel R)$是输入 R 和认证密钥 ID_k的 Hash 方程。然后阅读器向数据库申请获得所有射频标签的认证密钥。阅读器根据接收的 R 和存储在后台数据库中所有密钥的 ID_j计算 Hash 方程。如果 Hash 方程值与射频标签发送的 Hash 方程值匹配，则阅读器可以识别出该射频标签的密钥 ID_k并发送给射频标签。因为每次访问时，射频标签的输出改变了，因此该方法避免了被跟踪的缺点；但是由于被授权的阅读器识别一个射频标签，就需要搜索和计算所有标签的认证密钥 ID_k，因此该方法不适合大量射频标签和应用场景。随机化 Hash – Lock 认证协议流程如图 4-5 所示。

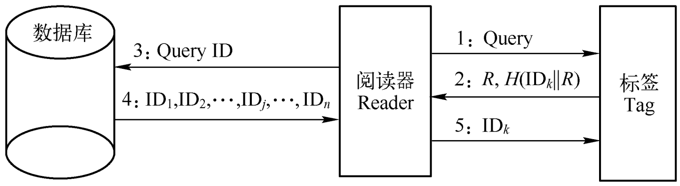


图 4-5　随机化 Hash – Lock 协议认证流程

随机化 Hash – Lock 认证协议的具体执行过程如下。

① 阅读器向标签发送 Query 请求。

② 标签生成一个随机数 R，并计算 $H(ID_k \parallel R)$，其中 ID_k是标签的认证密钥，标签将$(R, H(ID_k \parallel R))$发送给阅读器。

③ 阅读器向后台数据库提出获得所有标签认证密钥的 $ID_1, ID_2, \cdots, ID_j, \cdots, ID_n$的请求。

④ 后台数据库将自身存储的所有标签的认证密钥 $ID_1, ID_2, \cdots, ID_j, \cdots, ID_n$发送给阅读器，其中 n 为存储的标签数量。

⑤ 阅读器检查是否有某个 $ID_j(1 \leqslant j \leqslant n)$使得 $H(ID_j \parallel R) = H(ID_k \parallel R)$成立，如果有，则认证通过，并将 ID_j发送给标签；如果没有，则认证失败。

⑥ 标签验证 ID_j与 ID_k是否相同，如果相同，则认证通过。

随机化 Hash – Lock 认证协议中，认证通过后，标签密钥 ID_k仍然以明文形式进行传送，因此攻击者可以对标签进行有效的追踪。一旦获得了 ID_k，攻击者就可以实现对标签的假冒。同时，该认证协议也无法抵抗重放攻击，则随机化 Hash – Lock 认证协议的安全性也很低，且能够支持的标签数量有限，因此它在 RFID 系统中难以应用。

3. Hash 链认证机制

Hash 链认证协议是基于共享秘密的询问 – 应答协议。在系统运行前，标签 i 和后台数据库首先要共享一个初始密钥 $S_{i,0}$，标签和阅读器之间执行第 j 次（从 0 开始计数）Hash 链认证的过程如图 4-6 所示，具体如下。

① 阅读器向标签发送 Query 请求。

② 标签使用当前的密钥 $S_{i,j}$计算 $a_{i,j} = G(S_{i,j})$，并更新其密钥值为 $S_{i,j+1} = H(S_{i,j})$，同时将 $a_{i,j}$发送给阅读器。

③ 阅读器将 $a_{i,j}$转发给后台数据库。

④ 后台数据库维持一对（$ID_k, S_{k,0}$）的列表，针对所有的标签数据项查找并计算是否存在某个标签 k（标识为 ID_k）的密钥 $S_{k,0}$使得 $a_{k,j} = G(H^j(S_{k,0})) = a_{i,j}$，$G$ 也为单向函数。如果

有则认证通过，并将 ID_k（实际上就是 ID_i）发送给标签，否则认证失败。

从以上认证过程中可以看出，Hash 链协议是一个单项认证协议，它只实现阅读器对标签的认证，而没有实现标签对阅读器的认证。G 是单向函数，因此敌人能获得标签输出 $a_{i,j}$，但是不能从 $a_{i,j}$ 获得 $S_{i,j}$。G 输出随机值，敌人能观测到标签输出，但不能把 $a_{i,j}$ 和 $a_{i,j+1}$ 联系起来。H 也是单向方程，敌人能篡改标签并获得标签的密钥值，但不能从 $S_{i,j+1}$ 获得 $S_{i,j}$。该算法的优势很明显，但是有太多的计算和比较。为了识别一个 ID，后台服务器不得不计算 ID 列表中的每一个 ID。假设有 N 个已知的标签 ID 在数据库中，数据库不得不进行 N 次 ID 搜索，$2N$ 次 Hash 方程计算和 N 次比较。计算机处理负载随着 ID 列表长度呈线性增加，因此，该方法也不适合大量射频标签的情况。

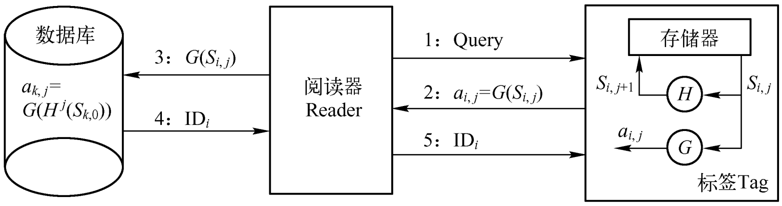


图 4-6　Hash 链协议的认证流程

4.2.2　RFID 分布式询问 – 应答认证机制

2005 年，Rhee 等人提出了分布式询问 – 应答认证协议，该协议是典型的询问 – 应答双向认证协议，其协议流程如图 4-7 所示。

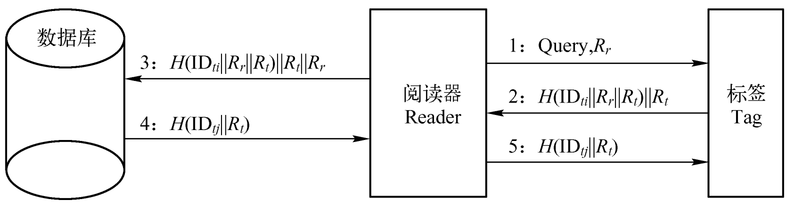


图 4-7　分布式询问 – 应答协议的认证流程

分布式询问 – 应答协议的执行过程如下。

① 阅读器生成一个随机数 R_r，并向标签发送认证请求：Query，R_r。

② 标签生成一随机数 R_t，并计算并发送 $H(ID_{ti} \| R_r \| R_t) \| R_t$ 给阅读器。

③ 阅读器将接收到的 $H(ID_{ti} \| R_r \| R_t) \| R_t$ 与随机数 R_r 相连接，并发送给后台数据库。

④ 后台数据库检查是否有某个 ID_{tj}，使得 $H(ID_{tj} \| R_r \| R_t) = H(ID_{ti} \| R_r \| R_t)$ 成立，如果有，则将 $H(ID_{tj} \| R_t)$ 发送给阅读器。

⑤ 阅读器接收到 $H(ID_{tj} \| R_t)$ 后，便完成了对标签的认证，并将 $H(ID_{tj} \| R_t)$ 转发给标签。

⑥ 标签验证 $H(ID_{tj} \| R_t) = H(ID_{ti} \| R_t)$ 是否成立，若成立，则标签完成对阅读器的认证。

到目前为止，还没有发现分布式询问 – 应答协议存在明显的安全漏洞或缺陷。但是该

协议中，执行一次标签认证，需要标签运行两次杂凑算法。标签的电路中自然也需要集成随机数产生器和杂凑函数单元，因此成本相对较高，不适合对成本要求较高的 RFID 系统。

4.2.3　RFID 轻量级安全认证

从前文的介绍中可以看出，基于 Hash 函数的 RFID 认证协议设计或多或少都需要标签进行较为复杂的运算，这对于降低标签成本仍然不利，为此很多更加简单的认证协议被提出来，这里主要介绍两种基于位运算操作的安全认证协议，分别是：轻量级强认证强完整性协议（Strong Authentication and Strong Integrity，SASI）和两消息互认证协议（Two - Message Mutual Authentication Protocol，T2MAP）。

1. SASI 认证协议

SASI 认证协议中，阅读器和标签之间共享密钥 K_1，K_2 和假名 IDS，其中假名 IDS 在认证的过程中用来代替标签的真实 ID，且在认证过程中不断更新。图 4-8 描述了 SASI 的认证过程，具体过程如下。

① 阅读器向标签发送 Query 认证请求。

② 标签发送假名 IDS 给阅读器。

③ 阅读器根据标签的假名，找到对应的密钥 K_1、K_2，并产生随机数 n_1、n_2，同时根据图中给出的计算方法计算 A，B，K_1，K_2 和 C，同时将 $A \parallel B \parallel C$ 发送给标签。

④ 标签根据 A、B 计算 n_1、n_2，再计算出 $\overline{K_1}$、$\overline{K_2}$，并得到 C，比较接收到的 C 与计算得到的 C 是否一致，如果一致则完成对阅读器的认证，若认证成功则将 D 发送给阅读器。

⑤ 阅读器将收到的 D 与自己计算的 D 进行比较，如果相一致，则完成阅读器对标签的认证，反之，认证失败。

⑥ 相互认证后，标签和阅读器分别根据 $\mathrm{IDS}=(\mathrm{IDS}+\mathrm{ID})\oplus(n_2\oplus\overline{K_1})$，$K_1=\overline{K_1}$ 和 $K_2=\overline{K_2}$ 对密钥和假名进行更新。

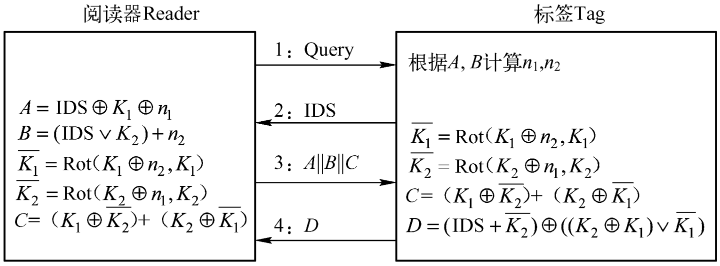


图 4-8　SASI 的认证过程

SASI 使用简单的异或、移位和与操作代替复杂的密码算法，简化了标签的计算成本，标签硬件电路可以减少到 300 个左右逻辑门，完全满足 RFID 系统低成本的设计要求，同时具有较好的前向安全特性和双向认证功能。

2. T2MAP 认证协议

T2MAP 仅需要两条消息就可以实现双向认证。该协议规定在标签和阅读器的内存中保存相互对应的标签 ID 和密钥，图 4-9 描述了 T2MAP 的认证过程。

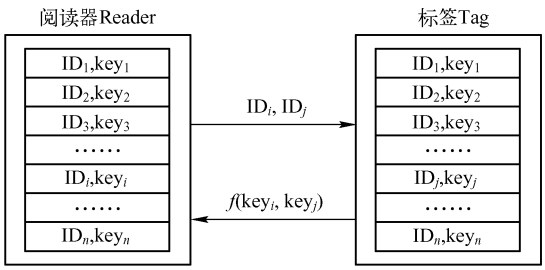


图 4-9　T2MAP 的认证过程

协议的执行过程如下。

① 阅读器向标签发送认证请求，阅读器从内存的 ID 序列中，随机选择两列 ID_i、ID_j发送给标签。

② 标签在自己的内存序列中查找是否有相同的 ID_i、ID_j，如果满足，则完成标签对阅读器的认证。

③ 标签将 ID_i、ID_j对应的密钥 key_1、key_2，经过简单的操作（如异或操作）后传送给阅读器。

④ 阅读器将收到的 $f(key_1, key_2)$ 与自己计算得到的 $f(key_1, key_2)$ 比较，如果一致则完成阅读器对标签的认证。

T2MAP 是目前最简单的认证协议，不需要任何加密电路，但协议的安全性也较低，一旦某一个标签被攻击者捕获并破解，使得存储的信息泄露，则攻击者便可以欺骗阅读器，也可以伪装成阅读器欺骗合法标签；同时由于标签资源有限，能够存储的标签 ID 和密钥也非常有限，因此采用该协议的 RFID 系统不能使用太多的标签。

4.2.4　一种基于 PUF 的 RFID 认证协议

基于 PUF 的认证技术主要完成的是设备的实体认证。下面将 Lars Kulseng 在 “Lightweight Mutual Authentication and Ownership Transfer for RFID Systems” 一文中的 RFID 相互认证协议介绍给大家。该协议的主要过程包括初始化和认证两个阶段，具体描述如下。

1. 前提考虑

- 系统中标签的 ID 应该受到保护。保护的方法是利用一个称为 IDS 的假名来代表标签，该随机数在每轮的认证中都会更新。
- 访问序号应该受到保护。具体的方法是利用随机数通过异或运算实现对访问序号的有效加密。
- 认证过程需要执行多次。标签在使用过程中需要进行多次认证。在每一次的认证过程中，访问序号都会被更新，目的就是为了使阅读器和标签能够直接进行多次相互

认证。

2. 初始化阶段

在初始化阶段，阅读器和标签共享如下信息。

LFSR：线性反馈移位寄存器，标签和阅读器内部都有的硬件电路。

ID：标签的 ID 值。

IDS：标签的一个假名，该假名每次认证都会更新，阅读器内会存放标签的相关元组信息，而假名的作用就是用作标签相关元组信息的索引。

G_n：阅读器发给标签的访问序号，n 代表相互认证的轮次，初始值为 1。

F：$[1,q] \rightarrow [1,q]$：随机置换函数 F，该函数能够生成一个 $[1,q]$ 范围内的随机数，这里 $\log q$ 的值为标签 ID 的比特长度；F 在此处充当随机数发生器，该函数可以是公开的，为了能够适应 RFID 系统，该函数在硬件平台上实现时要尽量高效并且开销小。

除了以上信息是标签和阅读器共享的以外，阅读器还要存储如下信息。

G_{n+1}：下一轮认证过程中阅读器发送给标签的访问序号，该值也用作本轮认证过程中标签发送给阅读器的访问序号。

另外，标签自身还要实现一个函数 P，描述如下。

P：$[1,q] \rightarrow [1,q]$：P 函数和 F 函数类似，都是生成一个 $[1,q]$ 之间的随机数，但 P 函数的特点是每个标签的 P 函数都不同。在认证过程中 P 函数被用作阅读器对标签的认证。另外一个 P 函数的属性就是在被攻击者攻击的条件下，P 函数是无法被重新构建的。实际上 P 函数就是标签内部的 PUF 单元。

3. 认证过程

协议的认证过程如图 4-10 所示。

① 阅读器连续地向标签广播 Req 请求。

② 标签接收到 Req 请求后，会向阅读器返回自身的假名 IDS，由于采用的是假名，不会泄露标签的真实 ID。

③ 阅读器根据接收的 IDS 查找内部存储的与该标签相关的元组，找到后发送 $\mathrm{ID} \oplus G_n$ 信息给标签。本次通信过程中，由于 ID 信息受到 G_n 的保护，不会泄露出去，因此想利用 ID 信息来跟踪标签是不可能的。标签接收到该信息后，根据自身的 ID 和内部存储的 G_n 来验证该信息的正确性，如果正确，则说明标签是合法的，因为只有合法的标签才有 ID 和 G_n 的值。

④ 验证完阅读器后，标签会计算生成两个访问序号，分别是 G_{n+1} 和 G_{n+2}，具体的计算方法如下：

$$G_{n+1} = P(G_n)$$

$$G_{n+2} = P(G_{n+1}) = P^2(G_n)$$

接下来，标签利用 G_n 和函数 F 计算 K_n 和 K'_n，计算方法如下：

$$K_n = F(G_n)$$

$$K'_n = F(K_n) = F^2(G_n)$$

然后，标签发送 $G_{n+1} \oplus K_n$，$G_{n+2} \oplus K'_n$ 给阅读器，接收到这两个信息后，阅读器就可以

利用 G_n 计算 K_n，并从 $G_{n+1} \oplus K_n$ 中恢复 G_{n+1}。如果恢复出来的 G_{n+1} 与阅读器存储的 G_{n+1} 是一样的，阅读器就可以认为标签是合法的，因为只有该标签的 P 函数才能生成 G_{n+1}。一旦标签认证完成，阅读器就可以从接收到的 $G_{n+2} \oplus K'_n$ 信息中提取出 G_{n+2}。G_{n+2} 可以用于下一轮的认证过程中。

最后，标签和阅读器可以协商更新 IDS。假定用 IDS_{old} 代表刚刚在第二步认证过程中使用的假名，用 IDS_{new} 代表新的假名，则新的假名生成办法如下，其中 LFSR 是标签和阅读器内部的线性反馈移位寄存器。

$$IDS_{new} = LFSR(IDS_{old} \oplus G_n)$$

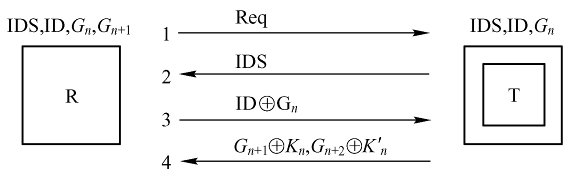


图 4-10　基于 PUF 的 RFID 相互认证协议的认证过程

该协议能够很好地抵抗窃听攻击、针对标签的物理攻击、重放攻击等多种攻击行为，是一种典型的 PUF 在 RFID 系统中的认证应用方法。

4.3　传感器网络认证技术

针对数据机密性、数据完整性、信息认证及数据新鲜性等安全特性，A. Perrig 等提出了传感器网络安全协议 SPINS，其中包含两个子协议：SNEP 和 uTESLA。SNEP 提供了基本的安全机制——数据机密性、双方数据鉴别和数据新鲜度，uTESLA 是传感器网络广播认证协议。本节重点介绍 SNEP 和 uTESLA 传感器网络认证协议，同时还将介绍非常具有应用前景的面向移动传感器网络的基于身份加密的认证协议和基于 PUF 的传感器网络认证协议。

4.3.1　SNEP 网络安全加密协议

SNEP 是为传感器网络量身打造的，具有低通信开销的，能够实现数据机密性、完整性、保证新鲜度的简单高效的安全协议。SNEP 只描述了安全实施的协议过程，并没有规定实际使用的算法，具体算法在具体应用时予以考虑。SNEP 采用共享主密钥的安全引导模型，假设每个节点都和基站之间共享一对主密钥，其他密钥都是从主密钥中衍生出来的。SNEP 的各种安全机制都是通过信任基站完成的。SNEP 具有以下特点。

- 语义安全：在每个信息之后，计数值都被增加，同样的信息在不同时刻被加密后产生的密文不同。
- 数据认证：要是 MAC 校验正确，消息接收者就可以确信消息发送者的身份。
- 重放保护：MAC 计数的方法，防止了消息的重放。
- 新鲜度保证：要是信息验证无误，接收者便知道该信息一定是在计数值较小的信息之后发送的信息。

- 低通信开销：计数器的状态保持在每一个节点上，不需要在节点间发送，降低了通信开销。

为了方便后面描述，表 4-3 给出了需要注意的表达式及符号以描述该协议。

表 4-3　需要注意的表达式及符号

表达式及符号	解　释
A,B	传感器节点
N_A	节点 A 产生的随机数（Nonce）
$M_1 \parallel M_2$	表示信息 M_1，M_2 串联
k_{AB}	表示 A，B 之间共享的对称密钥
$\{M\}_{k_{AB}}$	使用共享密钥 k_{AB} 对消息 M 进行加密
$\{M\}_{(k_{AB},\mathrm{IV})}$	使用共享密钥 k_{AB} 对消息 M 进行加密，其中 IV 为被使用在加密模型中的初始化向量

1. SNEP 的完整性和点到点的认证

语义安全是 SNEP 的一个优点，不但能够保证数据的机密性，还能保证即便是攻击者知道一些明文密文对的情况下也不能获得明文的任何信息。加密机制中实现语义安全有多种模式，其中计数器模式就是重要的一种方法。计数器模式可以实现语义安全的主要原因就在于每个数据包的密文与其加密时的计数器的值有关。同样，在 CTR 模式中，通信双方也共享一个计数器，计数器的值作为每次通信加密的初始化向量。这样每次通信时的计数器值不同，相同的明文也必定产生不同的密文。SNEP 正是结合这些机制才具有了语义安全的特性。SNEP 中的加密数据遵循以下格式，其中 D 是数据，加密密钥是 k_{encr}，C 是计数器的值。

$$E=\{D\}_{(k_{\mathrm{encr}},C)}$$

SNEP 的消息完整性和点到点认证是通过消息认证码（MAC）协议实现的。消息认证码协议的认证公式如下：

$$M=\mathrm{MAC}(k_{\mathrm{mac}},C \parallel E)$$

A 发给 B 的完整信息为：

$$A \rightarrow B:\{D\}_{(k_{\mathrm{encr}},C)},\mathrm{MAC}(k_{\mathrm{mac}},C \parallel \{D\}_{(k_{\mathrm{encr}},C)})$$

这里的 k_{mac} 表示消息认证码算法密钥，$C \parallel E$ 为计数器的值 C 和密文 E 的串联，表示消息认证时对计算器和密文一起运算。消息认证的内容可以是明文也可以是密文，SNEP 采用密文的方法，这样可以加快接收节点对消息的认证速度，降低系统对 DoS 攻击的敏感性。另外，无线传感器网络是多跳网络，而采用逐跳认证方式只能选择密文认证，因为中间节点没有端到端的通信密钥，所以不能对加密的数据进行解密。

k_{mac} 和 k_{encr} 是从主密钥 k_{master} 按照相同的计算方法推导出来的。SNEP 没有定义推导算法，在实现时，按照一定的规则生成。例如，可以直接用单向生成函数 H 来生成 k_{mac} 和 k_{encr}：

$$k_{\mathrm{encr}}=H^{(1)}(k_{\mathrm{master}})$$

$$k_{\mathrm{mac}}=H^{(2)}(k_{\mathrm{master}})$$

2. SNEP 的新鲜性验证

SNEP 通过 CRT 模式支持数据通信的弱新鲜性，所谓弱新鲜性就是指一种单向的新鲜性

认证。假设节点 A 给节点 B 连续发送 10 个请求数据包：

$$A \to B: \{R_{A1}\}_{(k_{\text{encr}}, C_1)}, \text{MAC}(k_{\text{mac}}, C_1 \parallel \{R_{A1}\}_{(k_{\text{encr}}, C_1)})$$

$$A \to B: \{R_{A2}\}_{(k_{\text{encr}}, C_2)}, \text{MAC}(k_{\text{mac}}, C_2 \parallel \{R_{A2}\}_{(k_{\text{encr}}, C_2)})$$

……

$$A \to B: \{R_{A10}\}_{(k_{\text{encr}}, C_{10})}, \text{MAC}(k_{\text{mac}}, C_{10} \parallel \{R_{A10}\}_{(k_{\text{encr}}, C_{10})})$$

通过 B 的计数器值能够知道这 10 个请求是顺序从 A 发过来的，得到这 10 个请求后，B 将其交给上层进行处理，并回复给 A 10 个应答消息 RSP：

$$B \to A: \{\text{RSP}_{A1}\}_{(k_{\text{encr}}, C'_1)}, \text{MAC}(k_{\text{mac}}, C'_1 \parallel \{\text{RSP}_{A1}\}_{(k_{\text{encr}}, C'_1)})$$

$$B \to A: \{\text{RSP}_{A2}\}_{(k_{\text{encr}}, C'_2)}, \text{MAC}(k_{\text{mac}}, C'_2 \parallel \{\text{RSP}_{A2}\}_{(k_{\text{encr}}, C'_2)})$$

……

$$B \to A: \{\text{RSP}_{A10}\}_{(k_{\text{encr}}, C'_{10})}, \text{MAC}(k_{\text{mac}}, C'_{10} \parallel \{\text{RSP}_{A10}\}_{(k_{\text{encr}}, C'_{10})})$$

A 同样可以通过计数器的值判断这 10 个响应是从 B 顺序发过来的，而且对响应的重放攻击都能够被有效抑制，实现认证的弱新鲜性。这种认证有一个缺点就是 A 不能够判断它收到的响应包 RSP_{A1}是不是针对它发出的 R_{A1} 请求包的回应。如果 A 收到的回复消息不是按照其请求包发送的顺序给出的，那么它将无法为每个应答做出正确的响应。为此 SNEP 定义了强新鲜性认证方法。

SNEP 强新鲜性认证使用 Nonce 机制，Nonce 是一个唯一标识当前状态的任何无关者都不能够预测的数，因此它通常是由随机数发生器产生的。SNEP 在其强认证过程中，在每个安全通信的请求数据包中增加 Nonce 字段，用来唯一标识请求包的身份。为了保证安全性 Nonce 要足够长，以避免被攻击者预测。例如，节点 A 在发送给节点 B 的消息中增加一个 Nonce：N_A，B 在对该消息进行应答时，让 N_A参与认证的计算，并返回给 A，这样 A 就能够通过响应包的认证确定该应答对应哪个强求数据包。通信过程如下：

$$A \to B: N_A, \{R_k\}_{(k_{\text{encr}}, C)}, \text{MAC}(k_{\text{mac}}, C \parallel \{R_k\}_{(k_{\text{encr}}, C)})$$

$$B \to A: \{\text{RSP}_k\}_{(k_{\text{encr}}, C')}, \text{MAC}(k_{\text{mac}}, N_A \parallel C' \parallel \{\text{RSP}_k\}_{(k_{\text{encr}}, C')})$$

虽然强认证能够提升认证的安全性，但也带来了通信开销和计算开销的增加。

3. SNEP 完成节点之间的通信

SPINS 的安全引导条件是在节点和基站之间共享密钥对 k_{master}。如果直接使用这样的安全引导条件，则所有节点间的通信都必须经过基站。对于以查询方式运行的传感器网络应用，这样的安全处理完全能够胜任，因为节点感知的信息基本上都要传输给基站，由基站进行处理和转发。但对于在某些簇内需要进行通信的节点来说，此时如果通信也需要通过基站进行转发，显然通信效率会十分低下。一种可选的解决方案就是通过基站为需要通信的两个节点建立临时的通信密钥。如节点 A、B 需要进行通信，由于初始时两个节点没有任何共享密码，所以它们可以因彼此都信任的基站来协商建立安全通道。假设节点 A、B 和基站 S 分别共享密钥 k_{AS}和 k_{BS}。安全通道建立过程如下：

$$A \to B: N_A, A$$

$$B \to S: N_A, N_B, A, B, \text{MAC}(k_{BS}, N_A \parallel N_B \parallel A \parallel B)$$

$$S \to A: \{\text{SK}_{AB}\}_{k_{AS}}, \text{MAC}(k_{AS}, N_A \parallel B \parallel \{\text{SK}_{AB}\}_{k_{AS}})$$

$$S \to B: \{\text{SK}_{AB}\}_{k_{BS}}, \text{MAC}(k_{BS}, N_B \parallel A \parallel \{\text{SK}_{AB}\}_{k_{BS}})$$

SK_{AB}是基站为节点 A、B 设定的临时通信密钥，N_A、N_B是强新鲜认证所需要的 Nonce 随机数。在密钥协商过程中必须使用强新鲜认证，以确保协商的顺利。临时通信密钥可以在用后丢弃，新一次通信开销可以重新申请。

4.3.2 uTESLA 广播消息认证协议

1. uTESLA 的基本思想

物联网中，用户往往关注的是事件，而不是具体的某个设备。因此无线传感器网络中，用户经常会广播查询命令，来对感兴趣的事件进行查询。针对这种广播数据包，节点必须要能够对其进行来源认证，否则很容易受到 DoS 广播攻击。广播包的认证方法和单播包的认证有着非常大的区别，单播包的认证只需要收发节点之间共享一个密钥就可以完成，而广播包的认证则通常需要使用全网共享的公共密钥来完成，但这种方法的安全性很差，当某一个节点泄露了密钥信息后，整个网络都会受到安全威胁。如果网络频繁地更新公共密钥，则会带来非常大的网络开销，也不现实。传统解决广播数据包认证的方法是采用非对称数字签名的机制来实现，但非对称的加解密算法对于传感器节点而言无疑是难以完成的。

TESLA 协议提供了一个高效的广播认证方法，但由于此协议并不是针对资源受限的网络提出的，因此也难以在传感器网络中应用。Adrian Perrig 等人在 TESLA 基础上设计了针对传感器网络的 uTESLA 广播认证协议。其主要思想是先广播一个通过密钥 kmac 认证的数据包，然后公布密钥 kmac。这样就保证了在密钥 kmac 公布之前，没有人能得到认证密钥的任何信息，也就没有办法在广播数据包正确认证之前伪造出正确的广播数据包。uTESLA 广播认证协议解决的主要问题如下。

① 共享秘密问题。认证广播协议的密钥和数据包都通过广播方式发送给所有的传感器节点，因此必须防止恶意节点同时伪造密钥和数据包，为此节点必须能够首先认证公布的密钥，进而利用密钥认证数据包。uTESLA 采用的是全网共享密钥生成算法的方法，而不是共享密钥池（如 TESLA）。

② 密钥生成算法的单向性问题。由于密钥发布包是明文广播，且全网节点共享密钥生成算法，因此 uTESLA 协议为了防止恶意节点根据已知密钥明文和密钥生成算法推测出新的认证密钥，使用单向散列函数来解决密钥生成问题。这样即使恶意节点拥有了算法和已经公开的密钥，仍然无法推算出下一个要公布的密钥。

③ 密钥发布包丢失问题。uTESLA 要求基站密钥池中存放的密钥不是相互独立的，而是经过单向密钥生成算法迭代运算产生出来的一串密钥。已知祖先密钥，可以利用单向生成函数获得所有的子孙密钥。这样即使个别密钥发布包丢失，节点仍然可以根据最新的密钥把它们推算出来。

④ 时间同步和密钥公布延迟问题。uTESLA 使用周期性公布密钥的方式，一段时间内使用相同的密钥。周期性地更新密钥要求基站与传感器节点之间要维持一个简单的同步，这样节点就可以根据当前的时间来判断公布的密钥是哪个时间段使用的密钥，然后对该时间段接收的广播数据包进行认证。当然，密钥的周期及密钥广播的延迟时间需要权衡，时间太短则网络通信开销大，时间太长则节点容易存储过多的数据包。

⑤ 密钥认证和初始化问题。节点对每个收到的密钥首先要确定其来自基站，而不是攻

击者伪造的。密钥生成算法的单向特性为密钥的确认提供了很好的手段。因为密钥是单向可推导的，所以已知前面获得的密钥合法，则可以验证新的密钥是否合法。这个认证过程要求初始第一密钥必须确认合法。这个初始确认是通过协议的初始化完成的。uTESLA 使用 SNEP 来进行初始认证密钥和同步时间的协商。

2. uTESLA 过程描述

uTESLA 的运行过程包括基站安全初始化、网络节点加入安全体系和节点完成数据包的广播认证三个过程。

基站一旦在目标区域内开始工作，首先生成密钥池，确定密钥同步时钟。密钥池的大小 N 和密钥同步周期 T 一般根据实际情况来确定。一旦密钥池中的密钥消耗完，则需要重新实施网络初始化和时间同步。

假设 $F(x)$ 是单向密钥生成函数，N 为密钥池的大小，K_N 为基站确定的初始密钥，对于任意的密钥 K_{i+1}，运用 $K_i = F(K_{i+1})$ 得到其子密钥。运行 N 次密钥生成过程，得到大小为 N 的密钥池：$F^{(0)}(K_N), F^{(1)}(K_N), F^{(2)}(K_N), \cdots, F^{(N)}(K_N)$，其中 $F^{(N)}(K_N)$ 对应 K_0。

基站生成广播认证的密钥后需要定义两个变量：同步间隔 T_{int} 和密钥发布延迟时间间隔 $d \times T_{\text{int}}$。同步间隔表示一个广播密钥的生存期，在一个同步周期内 $(i \times T_{\text{int}}, (i+1) \times T_{\text{int}})$ 基站发布的广播包使用相同的密钥 K_i。密钥发布延迟定义为同步周期的一个整倍数，并且要求至少大于基站和最远节点之间的一次包交换的时间，这样以保证最远节点收到一个广播数据包时，该数据包的认证密钥还没有公布出来。

基站完成广播安全初始化后，就开始接受节点的加入。每个节点通过 SNEP 与基站之间建立同步。假设节点 A 在 $(i \times T_{\text{int}}, (i+1) \times T_{\text{int}})$ 时间段内向基站 S 要求加入网络，则其加入的具体过程如下：

$$A \to S: (N_M \parallel R_A)$$
$$S \to A: (T_s \parallel K_i \parallel T_i \parallel T_{\text{int}} \parallel d), \text{MAC}(K_{AS,} N_M \parallel T_s \parallel K_i \parallel T_i \parallel T_{\text{int}} \parallel d)$$

其中 N_M 是一个随机数 Nonce，表示使用强新鲜性认证；R_A 为请求加入网络数据包；$K_{AS,}$ 是节点 A 与基站之间的认证密钥，通过预共享主密钥产生；T_s 为当前时间；K_i 为初始化密钥，T_i 是当前同步间隔的起始时间；T_{int} 是同步间隔；d 是密钥发布的延迟时间尺寸，单位为 T_{int}。经过这样一轮认证，节点将获得关于认证广播的所有信息。

节点加入网络可以发生在任何时刻，而认证广播的过程在基站初始化完成以后就可以进行了。图 4-11 是一个认证广播的实例，其中密钥公布延迟为两个时间单位。图 4-11 给出了基站连续五个密钥周期发送广播包 $P_1 \sim P_8$ 和公布 $K_{i-1} \sim K_{i+3}$ 的过程。节点在收到广播包 P_1, P_2 时，通过时间同步条件判断它们使用的广播认证密钥 K_i 还没有公布出来，此时将两个数据包保存起来。节点拥有基站的密钥公布时间表，因此它会在基站公布该密钥时查收这个密钥。收到密钥 K_i，节点首先计算 $F(K_i)$ 是否与 K_{i-1} 相同，如果相同 K_i 就是合法密钥，否则丢弃。收到合法密钥 K_i 以后，节点将根据时间标尺自动使用 K_i 认证在 $[T_i, T_i + T_{\text{int}})$ 时间段内收到的广播包 P_1 和 P_2。

假设节点没有收到 K_{i+1} 这个密钥，节点会把 P_3 包的认证推迟到接收下一个广播密钥 K_{i+2} 时。节点收到 K_{i+2} 以后，首先根据时间标尺知道这个密钥是认证 $[T_{i+2}, T_{i+2} + T_{\text{int}})$ 时间

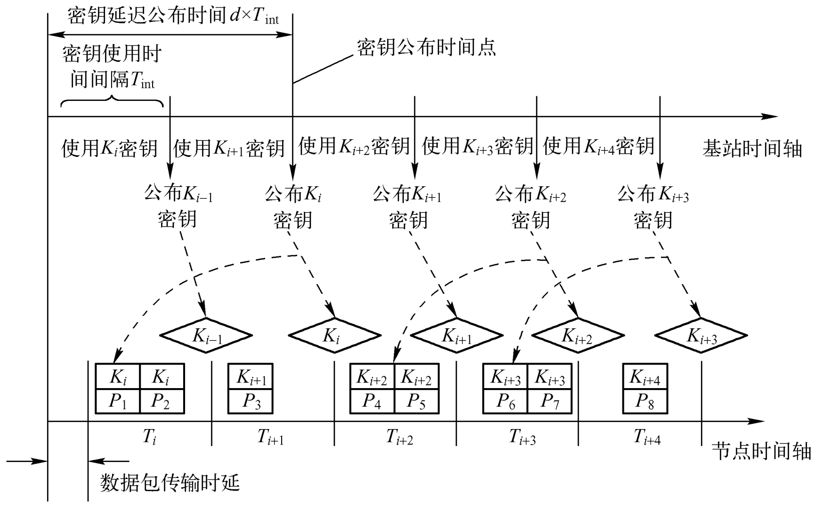


图 4-11　uTESLA 协议广播认证过程说明图

段的密钥。由于 K_{i+1} 没有收到，节点会通过判断 $F^{(2)}(K_{i+2})$ 是否等于 K_i 确定密钥的合法性。

uTESLA 没有考虑 DoS 攻击问题。恶意节点广播错误数据包，节点会将这些数据包保存起来等待验证，这样可能耗尽节点资源。至于如何抵抗 DoS 攻击，相关研究也提出了具体的方法，这里不详细分析。

3. 多级 uTESLA

如前所述，一个节点要加入网络，需要与基站使用 SNEP 完成密钥初始化和同步过程。这个过程是一个点对点的单播过程。对于规模较大的网络而言，网络边缘的节点要想完成广播协议的初始化常常会耗费非常大的网络资源，因此 uTESLA 适用的网络规模有限。Liu 等提出一种多级 uTESLA 协议。该协议首先引进了预定和广播初始化参数的方法；另外，该协议使用冗余传输机制和随机选择策略来完成密钥链发布任务，以提高网络对包丢失的容忍度和抗 DoS 攻击能力。

预定和广播初始化参数的方法可以避免节点加入网络时的 SNEP 安全认证过程和参数协商。它的基本思想是在节点布置以前，所有节点都已知网络的密钥链和相关参数。这样节点在部署到目标区域后，节点认证广播消息的开始时间到达后，就可以直接对基站广播消息进行认证了。这样的单纯机制并不能得到满意的效果。为了能够使网络持续比较长的时间，就必须要维护一个比较长的密钥链，或者使用比较长的密钥更新间隔。两种方法都有各自的问题。如果采用很长的密钥链，基站需要很大的存储空间来维护这些密钥；另外，如果两个连续的广播消息之间有很长一段时间，而密钥公布时间很短，则为了确定后一个广播包的认证密钥的有效性，需要进行很多次单向散列函数的运算，浪费节点的资源。如果使用较长的密钥更新时间间隔，若广播包的频率很高，则节点将会存储过多的数据包。

为了能够解决密钥环长度和安全生命期之间的矛盾，新的协议提出了多级密钥链的概念。其基本思想是建立多个密钥链，其中高级密钥链用来认证低级密钥链，而最低级的密钥链用来认证广播包。低级密钥链采用比较高的广播频率，而高级密钥链则以低级密钥链的生命周期为周期进行切换。这样整个密钥链的生命周期是各级密钥链生命周期的一个乘积，从

而大大提高了安全体系的生命周期。不同级之间使用不同的散列函数产生密钥链，并且低级密钥链的源密钥随机选取。但是这个模型带来一个问题，就是容忍丢包的性能没有原始 uTESLA 那么好。首先，低级密钥链是断续的，一个密钥链用完以后，下一个低级的密钥环是独立于前一个密钥链生成的。一旦密钥链的最后一个更新密钥数据包丢失，它将不能够通过后面的密钥推导出这个丢失的认证密钥。另外一个问题是原始的 uTESLA 只需要考虑普通数据的丢失问题，而本机制又引入了高级密钥链更新密钥消息丢失的问题。高级密钥链更新消息会增大节点认证数据包的延迟，因为节点不得不等到当前低级密钥链全部使用完毕后才能够得到下一个高级密钥链的更新消息，从而衍生出丢失的高级密钥。

对于第一个问题，Liu 等人提出低级密钥不随机选取，而是与高级密钥链的密钥相关。这样节点可以通过高级密钥链推导出每个密钥链最后一个公布的认证密钥。不过这种方法需要节点存放另外一个单向散列函数来完成这种推导过程。对于高级密钥链公布密钥消息丢失问题解决起来则比较困难。如果是偶然的通信信道干扰导致的消息损坏，可以通过带有容错机制的通信编码方式缓解。对于在转发过程中丢失的情况，一种办法是在一个高级密钥公布消息中除了公布当前最新的更新密钥以外，还公布其前面的几个密钥。只要不是连续丢失密钥公布消息，就能够获得前面的高级密钥。但增长的数据包在无线信道中更容易丢失。另一种更为保险的设计是在一个高级密钥公布消息的周期内多次广播该消息，这样不必增大数据包的长度，却同样可以大大降低公布消息的概率。

多级 uTESLA 虽然有很多好的特性，但是复杂度高，并且占用更多的节点和计算资源。可以针对目标应用系统选择使用或部分使用其中的安全处理机制。

4.3.3　基于身份标识加密的身份认证

1984 年，Shamir 提出了一种公钥可以为任意字符串的公钥加密体制，称为基于身份标识的加密算法（Identity - Based Encryption，IBE）。该算法的主要思想是加密的公钥不需要从公钥证书中获得，而是直接使用标识用户身份的字符串。最初提出的这种基于身份标识加密算法的动机是为了简化电子邮件系统中证书的管理。当 Alice 给 Bob 发送邮件时，她仅仅需要使用 Bob 的邮箱 bob@ company. com 作为公钥来加密邮件，从而省略了获取 Bob 公钥证书这一步骤。当 Bob 接收到加密后的邮件时，联系私钥生成中心（Private Key Generator，PKG），同时向 PKG 验证自己的身份，然后就能得到私钥，从而解密邮件。

然而在 Shamir 提出 IBE 算法后的很长一段时间内都没有找到合适的实现方法，直到 2001 年，可实用的 IBE 算法由 Boneh 等提出，该算法利用椭圆曲线双线性映射来实现。该方案的安全性建立在 CDH（Computational Diffie - Hellman）困难问题的一个变形之上，称之为 BDH（Bilinear Diffie - Hellman）问题。Boneh - Franlin 的 IBE 算法可参考文献“Identity - based encryption from the Weil pairing”，本章不再赘述。

基于 IBE 算法的加密签名一体化算法的主要思想是当节点有数据传输时，使用发送方的私钥对消息进行签名，再用接收方的公钥对消息进行加密并发送；接收方收到消息后，先解密消息，再利用解密的消息验证签名以验证消息是否由声明者发出。基于 IBE 算法的加密签名一体化算法可以将加密和认证结合起来，以较小的代价同时完成加密和认证，为网络密钥分配提供了新的加密手段。

基于身份的签名算法（Identity - Based Sign Cryption，IBS）由 Malone - Lee 于 2002 年在

文献“Identity－based signcryption”中第一次提出，Malone－Lee 在基于身份的密码系统基础上，利用签密的概念定义了基于身份的签密方案。然而该方案中的消息在签密密文中是可见的，这使得消息的机密性受到威胁；在此基础上，Malone－Lee 提出了一种改进的基于身份的加密算法，下面给出算法的一般形式和简要分析。

基于 IBE 的加密签名一体化算法包括相似的六个阶段，称为建立、提取、加密、签名、认证、解密。与一般的 IBE 加密算法相似的是，该算法先用发送方的私钥对消息进行签名，然后再用接收方的公钥加密签名的消息发送给接收方。接收方收到密文后，先解密出消息，再根据解密出的消息验证是否由声明的发送方发送以完成认证功能。算法的一般形式如下。

① 建立（setup）：给定安全参数 $k \in Z^+$，执行 G 输出 q，两个 q 阶群 G_1、G_2及双线性映射$\hat{e}:G_1 \times G_1 \rightarrow G_2$。随机选择 G_1生成元 $P \in G_1$，$H_1:\{0,1\}^{k_0+n} \rightarrow Z_q^*$，$H_2:G_2 \rightarrow \{0,1\}^{k_0+k_1+n}$。其中 k_0 为 G_1中成员的位数（bits），k_1 为身份标识 ID 的位数，n 为明文的长度。

② 提取（extract）：对于任意给定的字符串 $\mathrm{ID} \in \{0,1\}^*$，计算 $Q_{\mathrm{ID}} \in H_1(\mathrm{ID}) \in G_1^*$；计算私钥 $d_{\mathrm{ID}} = sQ_{\mathrm{ID}}$，其中 s 为主密钥。

③ 签名（sign）：发送者身份标识为 ID_s，明文 $m \in M$。计算 $Q_{\mathrm{IDS}} \in H_0(\mathrm{ID}_S) \in G_1^*$；随机选择数 $r \in Z_q^*$；计算 $X = rQ_{\mathrm{IDS}}$；计算 $h_1 \leftarrow H_1(X \| m)$和 $Z \leftarrow (r + h_1)d_{\mathrm{IDS}}$；签名为（$X,Z$），并向加密阶段提交四元组（$m,r,X,Z$）用于加密。

④ 加密（encrypt）：接收者身份标识为 ID_R，签名阶段的输出（m,r,X,Z）。计算 $Q_{\mathrm{IDS}} \in H_0(\mathrm{ID}_R) \in G_1^*$；计算 $w \leftarrow \hat{e}(rd_{\mathrm{IDS}}, Q_{\mathrm{IDR}})$和 $y \leftarrow H_2(w) \oplus (Z \| \mathrm{ID}_s \| m)$；密文为（$X,y$）。

⑤ 解密（decrypt）：接收者为 ID_R，给定密文 $C(X,y)$，私钥 d_{IDR}。计算 $w = \hat{e}(X, d_{\mathrm{IDR}})$和$(Z \| \mathrm{ID}_s \| m) \leftarrow y \oplus H_2(w)$得到消息 m 和签名（X,Z）；将得到的消息 m 和签名（X,Z）及发送者的身份标识 ID_s 作为解密阶段的输出。

⑥ 验证（verify）：为了验证消息是否由声明的发送方 ID_s 发送，计算 $Q_{\mathrm{IDS}} \in H_0(\mathrm{ID}_S) \in G_1^*$；计算 $h_1 \leftarrow H_1(X \| m)$；验证$\hat{e}(X,P) = \hat{e}(P_{\mathrm{pub}}, X + h_1 Q_{\mathrm{IDS}})$是否成立，如果成立则通过验证，否则终止算法。

4.3.4　基于 PUF 的延迟容忍传感器网络节点身份认证机制

前面提到的几种传感器网络认证机制，是以网络链路稳定且节点静止为前提的，而在物联网中，常常需要传感器节点部署在运动的物体上或部署在运动的媒体中（如水下），如对野生动物的生活习性进行监控或对水下信息进行采集等应用场景。这种具有节点运动特性和网络链路间歇性连通特性的网络称为延迟容忍移动传感器网络（Delay Tolerant Mobile Sensor Network，DTMSN）。在 DTMSN 中，传感器节点由于无法实时监控，因此很容易受到攻击者的捕获并破解，而又很难发现。一旦节点被破解，则攻击者便可以复制节点，实现对网络的克隆攻击。为了抵抗节点克隆攻击，文献“PUF－based node mutual authentication scheme for delay tolerant mobile sensor network”提出了基于 PUF 的 DTMSN 节点相互认证机制（Mutual Authentication scheme based on PUF，MAP），有效实现了对节点身份的合法性认证。

鉴于 DTMSN 的传感器节点资源受限且链路间歇连通，从实用的角度出发，DTMSN 相互认证机制必须满足以下需求。

- 抗克隆攻击：延迟容忍移动传感器网络中，节点通常长时间无法与其他节点进行通信，也无法彼此进行监控，因此为了防止攻击者利用克隆节点对网络实施攻击，节点在数据通信之前的相互认证中要能够识别克隆节点，抵抗节点克隆攻击。
- 高效性：网络中的传感器节点可能以较高的速度进行运动，因此节点间连接持续的时间通常非常短，这就需要节点间的认证过程尽量在较短的时间内完成，以便为数据传输保留时间。
- 一次一认证：由于节点间链路间歇连通，因此节点之间无法确保下一次连接到达时彼此间身份的合法性，为了提高网络的安全性，必须在每次进行数据传输前都要进行一次身份验证。

整个 MAP 认证过程分为两个阶段：初始化阶段和认证阶段。在具体介绍这两个过程之前，首先给出表 4-4 所列的表达式。

表 4-4　表达式

Id_i	节点 i 的身份标识
$\parallel$	连接符
$\rightarrow$	单播
$\Rightarrow$	广播
$\oplus$	异或操作
$H(s)$	对消息 s 进行 Hash 操作
$P_i(c)$	节点 i 的 PUF 单元对挑战 c 的响应
$\mathrm{MAC}(k,s)$	在密钥 k 下消息 s 的消息认证码

1. MAP 网络模型

如图 4-12 所示，DTMSN 网络中只有基站和普通的传感器节点，其中基站是网络的管理中心，同时负责节点信息的收集，网络中仅有一个基站；而传感器节点主要负责信息的采集和上传，节点间以“存储-携带-转发”的模式进行消息转发，网络中传感器节点的数量较多。网络中的每个传感器节点都有一个唯一的身份标识。基站静止不动，传感器节点可以随机运动。为了保证数据传输安全，节点间进行消息转发前首先要对彼此的身份进行认证。这里同时有如下假设。

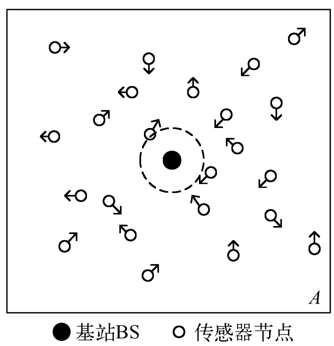


图 4-12　MAP 认证机制网络模型

① 网络中每个节点的芯片上都内嵌 PUF 单元，该 PUF 单元与芯片密不可分，任何分离 PUF 与芯片的操作或破解芯片的操作都将导致 PUF 单元的破坏。

② 基站在网络中被认为是可信的，并且不受资源和能量的限制，而传感器节点则资源和能量都有限。

③ 假设攻击者能力足够强大，不但能够侦听网络数据、向网络注入报文、实施重放攻击等操作，而且对于捕获的传感器节点能够实施破解，获取存储在芯片中的信息。但基站是安全的，不能被破解。

2. MAP 初始化过程

假设 DTMSN 网络中有 n 个传感器节点，其身份标识分别为 $\mathrm{Id}_1, \mathrm{Id}_2, \cdots, \mathrm{Id}_i, \cdots, \mathrm{Id}_n$。网络部署前，基站为每个节点生成一系列参数。首先，基站将为任意两个节点 i 和 j 随机生成一个共享的初始挑战 c_{ij}^0，并从节点 i 和 j 的 PUF 单元中获取得响应 $P_i(c_{ij}^0)$ 和 $P_j(c_{ij}^0)$。接下来基站选择单向散列函数 H（如 MD5）和消息认证码计算函数 MAC（如 HMAC），并且计算 $H(P_j(c_{ij}^0))$ 和 $H(P_i(c_{ij}^0))$。最后，基站将三元组 $<\mathrm{Id}_j, c_{ij}^0, H(P_j(c_{ij}^0))>$ 下载到节点 i 中，而将三元组 $<\mathrm{Id}_i, c_{ij}^0, H(P_i(c_{ij}^0))>$ 下载到节点 j 中，这样每个节点都存储有 $n-1$ 个其他节点的三元组信息。另外，HMAC 算法的实现程序也将被下载到每一个节点中。节点存储的三元组信息都只有节点自己知道，而基站则保留所有信息。

为了网络扩展需要，基站通常为每个节点选择大量的挑战，并获取响应的 PUF 进行存储，这些信息节点自身不知道，而当有新的节点加入网络中时，基站可以根据这些信息进行网络扩展。

3. MAP 认证过程

① 连接请求：延迟容忍移动传感器网络部署后，传感器节点如果有需要转发的消息，则通过周期性地广播 Hello 消息进行连接请求，同时也监听其他节点发起的连接请求。连接请求的 Hello 消息中包括节点的身份和一个随机生成的随机数 nonce，每次连接请求中的随机数都不同。假设网络中的任一节点 i 有数据进行转发，则 i 的广播消息如下：

$$i \Rightarrow * : \mathrm{Id}_i, \mathrm{nonce}_i$$

② 连接应答：假设节点 j 接收到了节点 i 的连接请求，并且也有意愿与 i 进行通信并建立连接，则节点 j 将遵循以下步骤对节点 i 进行应答。

- Step1：根据接收到的节点 i 的身份标识 Id_i 对自身存储的三元组进行查找（具体查找方法不在本书的研究范围之内），如果找到了三元组 $<\mathrm{Id}_i, c_{ij}^0, H(P_i(c_{ij}^0))>$ 则初步说明节点 i 是网络中的节点，否则放弃建立连接。
- Step2：节点 j 从查找到的三元组中提取共享初始挑战 c_{ij}^0，并计算 $c_{ij}^1 = H(c_{ij}^0)$（这里有 $c_{ij}^{k+1} = H(c_{ij}^k), k \geqslant 0$），同时将 c_{ij}^0、c_{ij}^1 输入自身的 PUF 单元获得响应 $P_j(c_{ij}^0)$，$P_j(c_{ij}^1)$。接下来节点 j 计算 $H(P_j(c_{ij}^1))$。
- Step3：节点 j 发出如下应答消息给节点 i，之后将 $P_j(c_{ij}^0)$、$P_j(c_{ij}^1)$ 和 $H(P_j(c_{ij}^1))$ 从其存储空间中擦除，防止信息泄露。

$$j \to i: \mathrm{Id}_j, P_j(c_{ij}^0) \oplus c_{ij}^0, H(P_j(c_{ij}^1)) \oplus c_{ij}^0, \mathrm{MAC}(c_{ij}^0,\ P_j(c_{ij}^0) \oplus c_{ij}^0 \parallel H(P_j(c_{ij}^1)) \oplus c_{ij}^0 \parallel \mathrm{nonce}_i)$$

③ 相互认证：节点 i 接收到节点 j 的应答消息后，首先对节点 j 的身份进行认证，具体认证步骤如下。

- Step1：节点 i 根据接收到的应答信息中的 Id_j 首先查找是否存在该节点的三元组 $<\mathrm{Id}_j, c_{ij}^0, H(P_j(c_{ij}^0))>$，如果不存在则放弃认证，如果存在则利用之前的随机数 nonce_i 对消息的完整性进行验证，通过消息完整性验证则说明消息没有被篡改，否则放弃验证过程。
- Step2：完整性验证通过后，节点 i 利用三元组中的 c_{ij}^0 从接收到的应答消息中提取

$P_j'(c_{ij}^0)=(P_j(c_{ij}^0)\oplus c_{ij}^0)'\oplus c_{ij}^0$，这里假设$(P_j(c_{ij}^0)\oplus c_{ij}^0)'$为接收的应答消息的第二部分。节点 i 计算 $H(P_j'(c_{ij}^0))$，并与三元组中的 $H(P_j(c_{ij}^0))$进行比较，如果相同，则说明节点 j 合法，否则非法，因为只有节点 j 能够生成正确的 $P_j(c_{ij}^0)$。

- Step3：在确认节点 j 合法后，节点 i 从应答消息中提取 $H(P_j(c_{ij}^1))$并存储。同时，节点 i 计算 $c_{ij}^1=H(c_{ij}^0)$，并将将 c_{ij}^0、c_{ij}^1输入自身的 PUF 单元获得响应 $P_i(c_{ij}^0)$、$P_i(c_{ij}^1)$。同时计算 $H(P_i(c_{ij}^1))$。完成以上操作节点 i 发送以下确认消息给节点 j。

$$i\rightarrow j:\mathrm{Id}_i,P_i(c_{ij}^0)\oplus c_{ij}^0,H(P_i(c_{ij}^1))\oplus c_{ij}^0,\mathrm{MAC}(c_{ij}^0,\ P_i(c_{ij}^0)\oplus c_{ij}^0\parallel H(P_i(c_{ij}^1))\oplus c_{ij}^0\parallel \mathrm{nonce}_i)$$

- Step4：节点 j 在接收到节点 i 的确认消息后，同样采取 Step1 和 Step2 中的操作来确认节点 i 的身份合法性，如果非法则放弃认证，如果合法则从节点 i 的确认消息中提取 $H(P_i(c_{ij}^1))$并存储。

④ 参数更新：从安全角度考虑，节点 j 和 i 在确认对方身份合法后，需要将存储的对方的三元组信息进行更新，以用于下一次身份验证。参数更新的过程比较简单，节点 i 和 j 将原有三元组中的 c_{ij}^0用 c_{ij}^1进行替换，同时将 $H(P_i(c_{ij}^0))$换成 $H(P_i(c_{ij}^1))$。当节点 j 和 i 下一次相遇时，则利用新的三元组信息进行相互认证。参数的更新能够有效提高网络的安全性，避免重放攻击。然而为了避免节点受到攻击，参数更新通常是在正常的数据交换后进行的。图 4-13（a）和图 4-13（b）中给出了第 $k-1$ 次和第 k 次认证的过程。

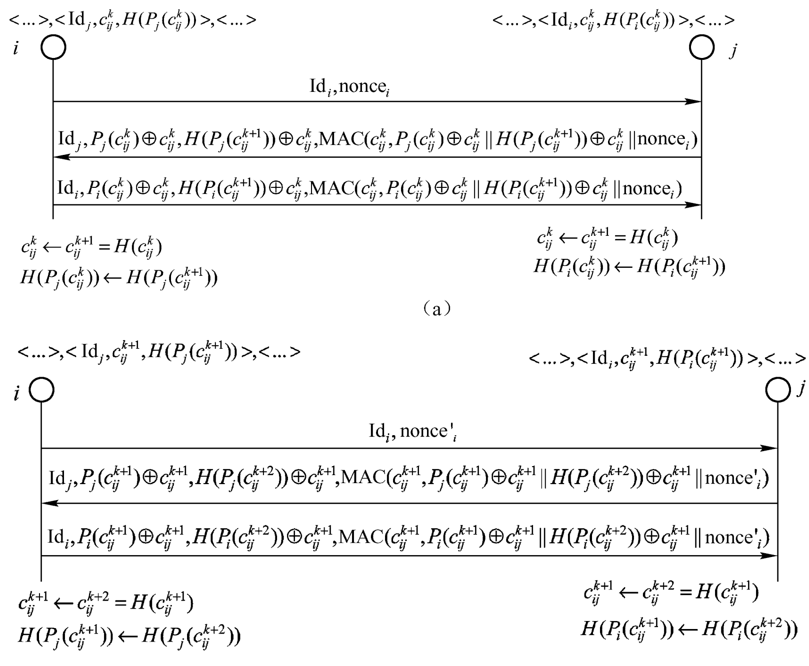


图 4-13　MAP 认证过程

物理不可克隆功能是 MAP 机制安全的基础。在该机制中，PUF 单元与传感器节点的芯片密不可分，任何破坏芯片的操作都将导致 PUF 单元结构的损坏，因此攻击者也无法探测 PUF 单元与芯片的内部通信数据。由于有关 PUF 的信息攻击者是无法探测到的，因此在

MAP 机制中，即便攻击者获取了存储在节点内部的三元组信息并且复制了恶意节点，但是由于 PUF 无法复制，则在与其他节点相互认证过程中是无法被其他节点认证通过的，因此 MAP 机制能够有效抵抗节点的克隆攻击。

参考文献

[1] 杨光．物联网安全威胁与措施清华大学学报．2011，51（10）：1335－1340.

[2] 王毅，镇维，等．物联网技术及应用．北京：国防工业出版社，2011.

[3] 朱磊．无线传感器网络安全认证若干关键技术研究［D］. 郑州：解放军信息工程大学，2010：22－23.

[4] BauerK，Lee H. A Distributed Authentieation Scheme for a Wireless Sensing System. In：Proceedings of 2nd International workshop on Networked Sensing Systems（INSS2005），2005：210－215.

[5] Wong K H M，Zheng Y，Cao J，et al. A dynamic user authentication scheme for wireless sensor networks，In：Proceedings of the IEEE International Conference on Sensor Networks，Ubiquitous，and Trust－worthy Computing（SUTC2006）. 2006：244－251.

[6] Perrig A，Szewezyk R，Tygar J，et al. SPINS：Security protocols for sensor networks［J］. ACM Wireless Network，2002，8（5）：521－534.

[7] Karlof C，Sastry N，Wagner D. Tinysec：a link layer security architecture for Wireless sensor networks. In：Proceedings of the 2nd international conference on Embedded networked sensor systems. ACM Press，2004：162－175.

[8] Li T，Wu H，Wang X，et al. Sensec Design，TR－I2R－vl. 1. InfoComm Secunty Department，Institute for Infocomm Research，2005.

[9] Luk M，Mezzour G，Perrig A. Minisec：a secure sensor network communication architecture . In：Proeeedings of the 6th intemational conference on Information Processing in sensor networks. ACM Press，2007：479－488.

[10] Guajardo J，Kumar S S，Schrijen G－J，et al. FPGA Intrinsic PUFs and Their Use for IP Protection［C］. In P. Paillier and I. Verbauwhede，editors，Cryptographic Hardware and Embedded Systems（CHES 2007），47（27），September 2007：63－80.

[11] Guajardo J，Kumar S S，Schrijen G－J，et al. Physical Unclonable Functions and Public Key Crypto for FPGA IP Protection［C］. In：International Conference on Field Programmable Logic and Applications（FPL 2007），August 27－30，2007：189－195.

[12] Holcomb D E，Burleson W P，Fu K，Initial SRAM State As A Fingerprint And Source of True Random Numbers For RFID Tags. Conference on RFID Security 07，July 11－13，2007.

[13] Gassend B，Clarke D，Dijk M，et al. Silicon Physical Random Functions［C］，In：Proceedings of the 9th ACM Conference on Computer and Communications Security（CCS’02），November 2002：148－160.

[14] Cortese Pier Francesco，Gemmiti Francesco，Palazzi Bernardo，et al. Efficient and Practical Authentication of PUF－Based RFID Tags in Supply Chains［C］，Program for the IEEE International Conference on RFID－Technology and Applications，Guangzhou，China，June 2010：182－188.

[15] Kulseng L，Yu Z，Wei Y，et al. Lightweight Mutual Authentication and Ownership Transfer for RFID Systems［C］，In：Proc. INFOCOM，2010：251－255.

[16] 孙立民，李建中，陈渝，等，无线传感器网络，北京：清华大学出版社，2003.

[17] Yang Kuiwu，Guo Yuanbo，Wei Dawei et al. ，PUF－based Node Mutual Authentication Scheme for Delay Tolerant Mobile Sensor Network，The 7th International Conference on Wireless Communications，Networking and Mobile Computing，Wuhan，China，2011. 09：1856－1859.

[18] Cortese Pier Francesco, Gemmiti Francesco, Palazzi Bernardo, et al. Efficient and Practical Authentication of PUF – Based RFID Tags in Supply Chains, Program for the IEEE International Conference on RFID – Technology and Applications, Guangzhou, China, June 2010: 182 – 188.

[19] Kulseng L, Yu Z, Wei Y, et al. Lightweight Mutual Authentication and Ownership Transfer for RFID Systems, In: Proc. INFOCOM, 2010: 251 – 255.

[20] Tuyls P, Batina L, RFID – tags for Anti – Counterfeiting, Topics in Cryptology CT – RSA, Lecture Notes in Computer Science, Vol. 3860, San Jose, CA, 2006: 115 – 131.

[21] Malone – Lee J. Identity – based signcryption. cryptology ePrint archive. http://eprint. iacr. org

[22] Boneh D, Franklin M. Identity – based encryption from the Weil pairing. Lecture Notes in Computer Science, 2001, 2139: 21 – 229.

第 5 章　感知层密钥管理技术

无线网络密钥管理技术的研究成果非常丰富，大多数研究成果都致力于改进算法的安全性，一般的研究思路包含替换协议中的某个部件，或者设计新的协议抵抗攻击。然而，这些协议事实上并不能适用于资源受限的物联网节点，因为在协议中通常使用非对称密码体制，需要公钥管理机构支持，同时计算量庞大，会使电池等能源消耗殆尽。因此，利用对称密码体制设计轻量级的认证密钥管理方案，就成为物联网感知层密钥管理研究需要解决的重要问题。本章针对物联网感知层密钥管理的需求及网络特点，重点分析讨论了 RFID 及无线传感器网络的相关密钥管理方法及协议，给出了相关的典型密钥管理机制及常见协议。

5.1　感知层密钥管理技术概述

密钥管理的作用是确保密钥的真实性和有效性，能够保证密钥难以被窃取。即使被窃取了，由于密钥有使用范围和时间上的限制，不会造成更大的损失。同时密钥的分配和更换对用户而言是透明的，用户不一定要亲自掌握密钥。设计良好的密钥管理方案可以有效地保证用户的数据安全。

5.1.1　RFID 密钥管理技术

按照密钥管理的对象可以将其分为以下几种。

用户与阅读器之间的密钥分配，这类密钥管理方案主要是对用户的合法性进行认证，通过认证后，则可信密钥管理服务器为用户分配某个阅读器和用户之间的临时通信密钥。

阅读器和标签之间的密钥分配，当阅读器需要读取某标签信息时，必须输入该标签的密码，保证了标签信息不会被非法读取。

用基于可信第三方的密钥管理协议来产生标签、阅读器或应用之间的会话密钥，在该方案中，网络中所有设备的密钥都由可信第三方进行管理，具体应用与阅读器之间是匿名的，有效地保护了阅读器的隐私性，并采用请求－相应动态标签认证，增加了标签和阅读器之间的互动。

5.1.2　传感器网络密钥管理技术

WSN 密钥管理最关键的安全需求是健壮性和自组织性。虽然机密性和完整性不容忽视，但只要密钥管理协议能安全地分发密钥，就基本能满足这两个需求。同样，数据新鲜性可以通过在每个数据包中添加随机数得到保证。但是，WSN 密钥管理的自组织需求却很难满足，需要密钥管理协议能适应 WSN 动态变化的网络环境。如果不考虑安全问题，WSN 具有自组织能力，节点能够自由建立连接，在出现故障（部分节点失效）时自动重构网络。如果密

钥管理协议给某些节点分发一定数目的通信密钥后，此需求可能难以满足。部分节点可能因缺乏适当的密钥而无法与某些节点建立安全连接。健壮性需求主要源自于 WSN 中部分节点可能被俘获而泄露存储的密钥信息。因为 WSN 通常部署在无人监护的检测区域，所以物理破坏是 WSN 需要面对的主要威胁之一。WSN 密钥管理协议应能容忍部分节点被俘获和分析。如果全网使用唯一密钥，一个节点被攻陷将导致整个网络失败。

与典型网络一样，WSN 密钥管理必须满足可用性（availability）、完整性（integrity）、机密性（confidentiality）、认证（authentication）和认可（non－reputation）等传统的安全需求。此外，根据 WSN 自身的特点，WSN 密钥管理还应满足如下一些性能评价指标。

- 可扩展性（scalability）。WSN 的节点规模少则十几个或几十个，多则成千上万。随着规模的扩大，密钥协商所需的计算、存储和通信开销都会随之增大，密钥管理方案和协议必须能够适应不同规模的 WSN。
- 有效性（efficiency）。网络节点的存储、处理和通信能力非常受限的情况必须充分考虑。具体而言，应考虑以下几个方面：存储复杂度（storage complexity），用于保存通信密钥的存储空间使用情况；计算复杂度（computation complexity），为生成通信密钥而必须进行的计算量情况；通信复杂度（communication complexity），在通信密钥生成过程中需要传送的信息量情况。
- 密钥连接性（key connectivity）。节点之间直接建立通信密钥的概率。保持足够高的密钥连接概率是 WSN 发挥其应有功能的必要条件。需要强调的是，WSN 节点几乎不可能与距离较远的其他节点直接通信，因此并不需要保证某一节点与其他所有的节点保持安全连接，仅需确保相邻节点之间保持较高的密钥连接。
- 抗毁性（resilience）。抵御节点受损的能力。也就是说，存储在节点或链路交换中的信息未给其他链路暴露任何安全方面的信息。抗毁性可表示为当部分节点受损后，未受损节点的密钥被暴露的概率。抗毁性越好，意味着链路受损就越低。

从 2003 年至今，WSN 密钥管理经历了一个研究高峰期，取得了许多成果。不同的方案和协议，其侧重点也有所不同。下面依据这些方案和协议的特点进行分类（见图 5-1）。

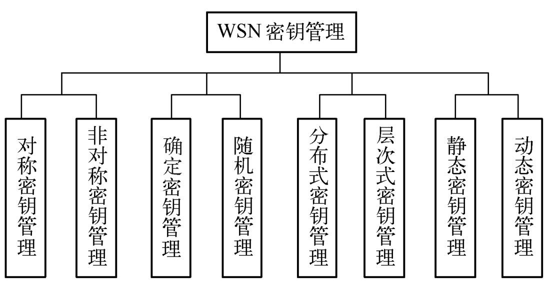


图 5-1　WSN 密钥管理机制分类

1. 对称密钥管理与非对称密钥管理

根据所使用的密码体制，WSN 密钥管理可分为对称密钥管理和非对称密钥管理两类。在对称密钥管理方面，通信双方使用相同的密钥和加密算法对数据进行加密、解密，对称密

钥管理具有密钥长度不长，计算、通信和存储开销相对较小等特点，比较适用于 WSN，目前是 WSN 密钥管理的主流研究方向。在非对称密钥管理方面，节点拥有不同的加密和解密密钥，一般都使用在计算意义上安全的加密算法。非对称密钥管理由于对节点的计算、存储、通信等能力要求比较高，曾一度被认为不适用于 WSN，但一些研究表明，非对称加密算法经过优化后能适用于 WSN。从安全的角度来看，非对称密码体制的安全强度在计算意义上要高于对称密码体制。

2. 确定密钥管理与随机密钥管理

根据节点的密钥分配方法不同，WSN 密钥管理可分为随机密钥管理与确定密钥管理。在随机密钥管理中，节点的密钥链（Key Ring）通过随机方式获取，如从一个大密钥池里随机选取一部分密钥，或从多个密钥空间里随机选取若干个密钥空间。而在确定密钥管理中，密钥链是以确定的方式获取的，如使用地理信息，或使用对称 BIBD（Balanced Incomplete Block Design）、对称多项式等。从连通概率的角度来看，随机密钥管理的密钥连通概率介于 0 和 1 之间，而确定密钥管理的连通概率总为 1。随机性密钥管理的优点是密钥分配简便，节点的部署方式不受限制；缺点是密钥的分配具有盲目性，节点可能存储一些无用的密钥而浪费存储空间。确定密钥管理的优点是密钥的分配具有较强的针对性，节点的存储空间利用较好，任意两个节点可以直接建立通信密钥；缺点是特殊的部署方式会降低灵活性，或密钥协商的计算和通信开销较大。

3. 分布式密钥管理和层次式密钥管理

根据网络结构，WSN 密钥管理可分为分布式密钥管理和层次式密钥管理两类。在分布式密钥管理中，节点具有相同的通信能力和计算能力。节点密钥的协商、更新通过使用节点预分配的密钥和相互协作来完成。而在层次式密钥管理里，节点被划分为若干簇，每一簇有一个能力较强的簇头（cluster head）来负责管理。普通节点的密钥分配、协商、更新等都通过簇头来完成。分布式密钥管理的特点是密钥协商通过相邻节点的相互协作来实现，具有较好的分布特性。层次式密钥管理的特点是对普通节点的计算、存储能力要求低，但簇头的受损将导致严重的安全威胁。

4. 静态密钥管理与动态密钥管理

根据节点在部署之后密钥是否更新，WSN 密钥管理可分为静态密钥管理和动态密钥管理两类。在静态密钥管理中，节点在部署前预分配一定数量的密钥，部署后通过协商形成通信密钥，通信密钥在整个网络运行期内不考虑密钥更新和撤销；而在动态密钥管理中，密钥的分配、协商、撤销操作周期性进行。静态密钥管理的特点是通信密钥无须频繁更新，不会导致更多的计算和通信开销，但不排除受损节点继续参与网络操作，若存在受损节点，则对网络具有安全威胁。动态密钥管理的特点是可以使节点通信密钥处于动态更新状态，攻击者很难通过俘获节点来获取实时的密钥信息，但密钥的动态分配、协商、更新和撤销操作将导致较大的通信和计算开销。

5.2　基于 HB 协议族的 RFID 密钥协商及管理技术

5.2.1　LPN 问题概述

LPN（Learning Paurity in the Presence of Noise）问题是为数不多的“矢量子集求和”困难问题，对计算量和存储量要求不高，适合 Tag 这样的设备，因此受到了设计者们的青睐。

假设 User(U)与 Computer(C)之间共享 k 比特密钥 x，U 想向 C 证明自己的身份，认证过程如下：C 产生一个随机的 k 比特矢量 a 发送给 U，U 收到后计算 $c=a\cdot x$ 响应给 C（其中 · 表示 GF(2)上的点乘），C 收到后检查，如果 $c=a\cdot x$ 则认证通过，反之不通过。在一轮认证中，C 接受一个假冒用户的概率是 1/2，重复 r 轮后，理论上 C 接受一个假冒用户的概率是 2^{-r}。很不幸，被动攻击者只要观察 $O(k)$ 次“挑战–响应”对，就可以通过高斯消元法解出共享密钥 x，进而伪装成 U。

引入噪声参数 $\eta\in(0,1/2)$，在 U 响应中加入一些错误的回答，这样被动攻击者就无法简单地利用高斯消元法得到密钥 x，这就是 LPN 问题，即噪声存在下的奇偶性问题。LPN 问题在不同的应用环境中有不同的描述，如 MDP 问题、Syndrome 译码问题等都是 LPN 问题的变型。下面用矩阵运算来定义 LPN 问题。

假设 D 是一个随机的 $q\times k$ 比特矩阵，x 是一个随机的 k 比特矢量，噪声参数 $\eta\in(0,1/2)$，v 是一个随机的 q 比特矢量，其汉明重量 $|v|\leqslant\eta q$，已知 D、η 及 $z=(D\cdot x)\oplus v$，找一个 k 比特矢量 x'满足 $|D\cdot x'\oplus v|\leqslant\eta q$。

LPN 问题已经被证明是 NP–Hard，同时要找到一个满足超过一半“挑战–响应”对的 x'也是 NP–Hard。

有研究结果表明，已知一个随机的 k 比特矢量 a，如果敌手可以用 k^{-c}的优势得到 $a\cdot x$ 值，那他就能解 LPN 问题。这就是一般地将某个协议规约到 LPN 问题的方法。后来又证明了 LPN 问题的伪随机性和 Log 一致性，并推测了 LPN 问题的困难性。

5.2.2　HB 协议

2001 年，Hopper 和 Blum 提出了一种用于 RFID 系统的轻量级认证协议：HB（Hopper and Blum）协议。该协议并没有如传统安全协议那样使用古典对称加密算法，而是使用了 LPN（Learning Parity with Noise）来提供安全性。后来有学者发现了该协议的安全缺陷，并提出了改进的方法，进而形成了 HB 协议族。

在 HB 协议族中，用到了以下符号和运算。

a,b：随机的 $n$bit 二进制向量。

x,x',y,y'：$n$bit 密钥向量。

v：噪声比特，$v=1$ 的概率 $P\in[0,0.5]$。

$a*x$：标量乘积。

$a\oplus x$：$n$bit 向量进行异或。

r：协议执行次数。

在 HB 协议中，标签和阅读器共享一个密钥，协议的流程为：

① 阅读器产生一个随机的挑战值 a，并发给标签；

② 标签计算 $z = a * x \oplus v$，并把 z 发给阅读器；

③ 阅读器收到后核对 $a * x$ 的值是否等于 z。

HB 协议一轮的流程如图 5-2 所示。该协议重复 r 次，如果阅读器核对标签发送来的 z 的失败次数不超过 $p * r$ 次，标签就通过了认证。由于 HB 协议是基于 LPN 的，所以它仅对被动攻击而言是安全的。对于简单的主动攻击，如攻击者伪装成阅读器，传送一个修改过的 a 给标签 n 次，就能推测到 x 的值。

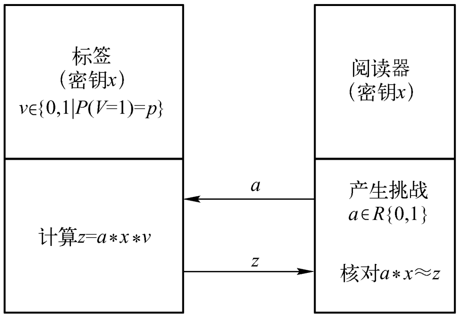


图 5-2　HB 协议流程

5.2.3　HB + 协议

Juels 和 Weis 对 HB 进行了修改，得到了 HB + 协议. 与 HB 协议不同的是，在 HB + 协议中加入了阅读器与标签共享的另一个 nb 密钥向量 y，并且从标签开始认证过程，即标签首先传送一个 nb 未知向量给阅读器。此外，HB + 协议还对计算公式进行了修改，新引进的密钥 y 和未知向量 b 的标量乘积与 HB 中的 z 需要进行异或运算。HB + 协议一轮的流程如图 5-3 所示。

Juels 和 Weis 展示了 HB + 对主动攻击是安全的，但是 Gilbert 证明了 HB + 协议对来自于伪装成一个有效标签阅读器的中间人攻击是不安全的。

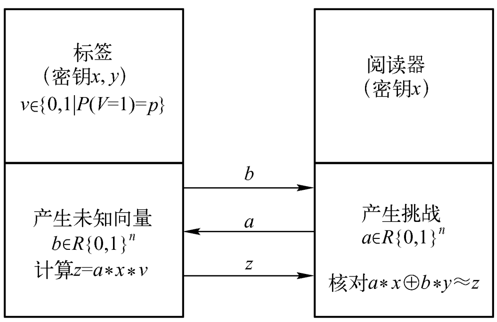


图 5-3　HB + 协议流程

5.2.4　HB ++ 协议

针对 Gilbert 对 HB + 的攻击，Bringer 提出了 HB ++ 协议，该协议中加入了阅读器与标

签共享的另两个 nb 密钥向量 x' 和 y'，并加入了 $z' = f(a) * \oplus x' f(b) * y' \oplus v'$，阅读器需要核对 z 和 z'。HB ++ 协议流程如图 5-4 所示。

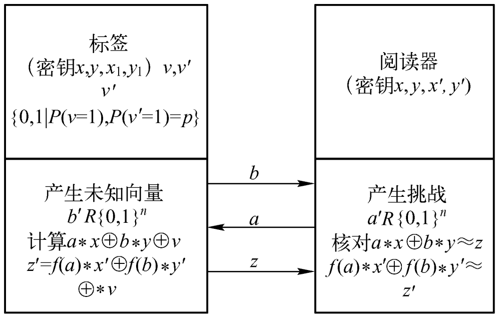


图 5-4　HB ++ 协议流程

5.3　传感器网络密钥分配及管理技术

5.3.1　预共享密钥机制

预共享密钥是最简单的一种密钥建立过程，主要包括以下几种方案。

1. 点对点密钥方案

网络中的每对节点之间都共享一个唯一的主密钥，以保证每对节点之间的通信都可以直接使用这个预共享密钥进行加密。该方案要求每个节点都存放与其他所有节点的共享密钥。如果网络中的节点数量为 N，则每个节点需要保存（$N-1$）个密钥。

该方案的优点：不依赖于汇聚节点，任意两个节点间的密钥是唯一的，因此节点被捕获不会泄露任何非直接建立的安全通信，可以达到最佳的安全性能。

该方案的缺点也显而易见：无法动态加入新的节点；网络规模巨大，节点为了存储密钥，需要大量的存储空间，由于节点存储器非常有限，因此该方案所能支持的网络规模较小。

2. 单一主密钥方案

网络中的所有传感器节点共享一个公共的主密钥 Master Key，任意一对节点之间的通信都直接使用这个主密钥进行加密，每个节点只需要保存该主密钥即可。

该方案的优点：非常容易实现，不依赖于汇聚节点；可任意加入新节点；由于每个节点只需要保存一个密钥，存储负载非常小，理论上能够支持的网络规模为无穷大。

该方案的缺点是致命的：方案基本不具备安全性，如果任意一个节点被捕获并取出主密钥，整个网络的安全性将不复存在。

3. 汇聚节点方案

网络中的每个普通节点与汇聚节点之间共享一个唯一的主密钥，这样每个节点只存储该主密钥即可，计算和存储压力全部集中在汇聚节点上。

该方案的优点：对普通节点资源和计算能力要求不高；支持的网络规模取决于汇聚节点的能力；汇聚节点还可以及时识别异构节点，并将其排除在网络之外。

该方案的主要缺点：过分依赖汇聚节点，如果普通节点被俘，会暴露与汇聚节点的共享密钥；如果汇聚节点被俘，则整个网络被攻破，因此要求汇聚节点被布置在物理上安全的位置。基于预共享密钥引导模型的方案虽然有很多不尽如人意的地方，但因其实现简单，所以在一些网络规模不大的应用中可以得到有效实施。

5.3.2　随机密钥分配机制

1. E－G 密钥分配协议

最安全的密钥分配协议是预先给每两个节点生成一个对偶密钥，把这些密钥保存在节点中，但是由于网络规模巨大，节点存储器非常受限，每个节点必须保存 $n-1$ 个密钥，可扩展性非常差，故只能用于小规模网络。Eschenauer、Gligor 引入随机图论，首先提出了基本的随机密钥预分配协议（简称 E－G 协议），旨在保证任意节点之间建立安全通道的前提下，尽量减少模型对节点资源的要求，这是当前传感器网络密钥协议的重点研究方向。

根据随机图理论，对于一个随机图 $G(n,P_r)$，n 是节点总数，如果要保证全图互连度 P 为一个很高的值（如 0.9999），每个节点无须确保和它的所有邻居节点建立安全链路，而只需要以不低于 P_{low} 的概率建立安全链路，通过其他多跳安全路径来建立与其他邻居节点的间接对偶密钥。

E－G 协议的基本思想是，一个比较大的密钥池，任何节点都拥有密钥池中的一部分密钥，只要节点之间拥有一对相同密钥就可以建立安全通道。如果节点存放密钥池的全部密钥，则基本随机密钥预分布模型就退化为点到点预共享模型。E－G 协议的建立过程由如下三阶段构成。

① 密钥初始化。首先随机产生一个非常大的密钥池 S，并为每一个密钥都分配一个编号 ID；在进行节点部署前，随机地从密钥池 S 中选取 m 个密钥分配给每个节点，这 m 个密钥称为节点的密钥环，$m \ll S$。

② 安全链路建立。当传感器节点被布置到目标区域后，节点开始进行密钥发现过程。各个节点通过广播自己密钥环中所有的密钥 ID，与邻居节点实现共享密钥发现，确定是否共享至少一个密钥；如果是，它们就可以利用其中编号最小的密钥作为它们的对偶密钥，称它们之间建立了安全链路。这种共享密钥发现可以通过每个节点以明文的形式广播自己拥有的全部密钥的编号实现。如果需要更高的安全性，可以采用挑战/应答方式：每一个节点用它的所有密码加密一个挑战随机数 a，并广播出去，而邻居节点用自己所有的密钥来尝试解密这些广播数据，如果能解密出 a，就确认双方共享此密钥，但是这样加解密的计算消耗比较大。

③ 间接对偶密钥建立。通过前一阶段，绝大多数节点通过安全链路组成了一个安全连接图。但前一阶段只能保证一定比例的邻居节点之间能建立对偶密钥，其他的邻居节点之间则可以通过这个安全连接图找到一条安全路径，通过它建立相互的对偶密钥。如果安全拓扑是连通的，则任何两个节点之间的安全路径总能被找到。

影响基本密钥预分配模型的安全连通性的因素有：密钥环尺寸 m、密钥池大小 $|S|$ 及两

者的比例、网络的部署密度（即网络通信连通度）、布置网络的目标区域状况。

$m/|S|$越大则相邻节点之间存在相同密钥的可能性越大。但 m 太大会导致节点资源占用过多，$|S|$太小或 $m/|S|$太大则会导致系统变得脆弱；当一定数量的节点被俘获以后，敌方人员将获得系统中绝大部分的密钥，导致系统秘密彻底暴露；网络部署度越高，则节点的邻居节点越多，能够发现具有相同密钥的概率就会比较大，整个图的安全连通概率也会比较高。对于网络布置区域，如果存在大量物理通信障碍，不连通的概率会增大。

根据随机图论，对于一个有 n 个节点的网络，且保证全网的互连度为 P_r，则每个节点建立安全链路的理想度数 d 如下式所示：

$$d=((n-1)/n)\times(\ln(n)-\ln(-\ln(c)))$$

d 与 n 及 P_r 的关系如图 5-5 所示。

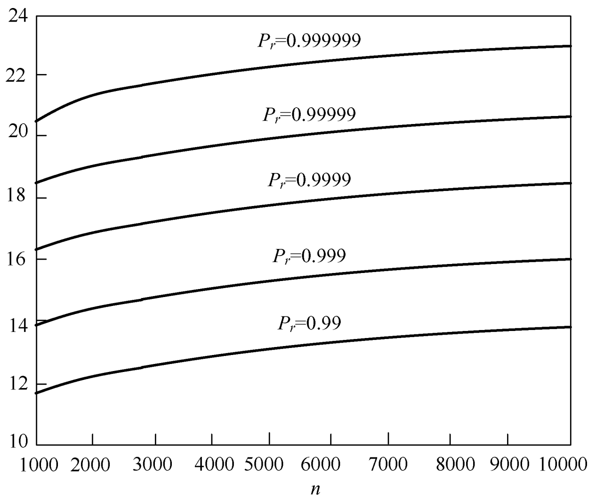


图 5-5　不同网络规模、互连度下的安全链路需求

对于一个给定密度的传感器网络，假设每个传感器节点在其通信半径内都有 n' 个邻居节点，为保证能建立 d 个安全链路，必须保证任意两个传感器节点之间以不低于概率‰ = dn' 的可能至少共享一个随机密钥。

假设密钥池总共有 S 个密钥，每个节点从中选择 m 个密钥，则任意两个节点至少共享一个密钥的概率为：

$$P_{\text{est}}=1-\frac{((S-m)!)^2}{S!(S-2m)!}\approx 1-\frac{\left(1-\frac{m}{S}\right)^{2(S-m+\frac{1}{2})}}{\left(1-\frac{2m}{S}\right)^{(S-2m+\frac{1}{2})}}$$

$$\text{当 } n \text{ 非常大时}, n!=\frac{\sqrt{2\pi}\,n^{n+\frac{1}{2}}}{e^n}$$

通过上述公式，我们可以根据已知的硬件条件选择合适大小的密钥池，保证整体网络达到自己期望的互连度值。n'是由传感器网络的特性决定的，P_r由设计者选择，可计算出每个节点的连通度 d 及任意两个节点间的最低连接概率 P_{low}；m 由节点硬件限制决定，选择最大的 S 满足 $P_{\text{est}}\geqslant P_{\text{low}}$。例如，假设有 10000 个节点的传感器网络，需要达到 $P_r=0.9999$ 的全

网互连度，则 $d=18.42$；假设 n' 为 40，则 P_{low} 为 0.4605；假设受硬件限制 m 为 200，选取最大的 S 满足 $P_{est} \geqslant P_{low}$，计算出 S 为 65017。

2. q－composite 随机密钥与分配模型

在基本模型中，任何两个邻居节点的密钥环中至少有一个公共的密钥。Chan H、Perrig A、Song D 在此基础上提出了 q－composite 随机密钥与分配模型，以一定代价有效地改进了 E－G 协议的安全性能。q－composite 模型主要对共享密钥发现有两点改进：要求节点间必须至少共享 q 个密钥才能建立安全链路；如果不少于 q，则采用它们之间所有的密钥来建立对偶密钥。该模型将这个公共密钥的个数要求提高到 q，提高 q 值可以提高系统的抵抗力，攻击网络的攻击难度和共享密钥个数 q 呈指数关系。但是要想安全网络中任意两点之间的安全连通度超过 q 的概率达到理想的概率值 p（预先设定），就必须缩小整个密钥池的大小，增加节点间共享密钥的交叠度。但密钥池太小会使敌人通包俘获少数几个节点获得很大的密钥空间。因此，寻找一个最佳的密钥池的大小是本模型的实施关键。

q－composite 随机密钥与分配模型和基本模型相似，只是要求相邻节点的公共密钥数要大于 q。在获得了所有共享密钥信息以后，如果两个节点之间的共享密钥数量超过 q，为 q' 个，就用所有 q' 个共享密钥生成一个密钥 K，$K=\text{hash}(k_1 \parallel k_2 \parallel \cdots \parallel k_{q'})$ 作为两个节点之间的共享主密钥。Hash 的自变量的密钥顺序是预先议定的规范，这样两个节点就能计算出相同的通信密钥。

任意两个节点之间恰好共享 i 个密钥的概率公式如下：

$$p(i)=\frac{\binom{S}{i}\binom{S-i}{2(m-i)}\binom{2(m-i)}{m-i}}{\binom{S}{m}^2}$$

两个节点之间能建立安全链路的概率为：

$$P_{est}=1-(P(0)+P(1)+\cdots+P(q-1))$$

对于给定的 m、P_{low}，我们选取最大的 S，使计算出的 $P_{est} \geqslant P_{low}$。$P_{low}=0.33$，$m=200$，E－G 协议与 $q=1$ 模式下，$S=10080$；$q=2$ 时，$S=33928$；$q=3$ 时，$S=19758$。很显然，q 越大，S 越小。

q－composite 模型相对于基本随机密钥预分布模型对节点被俘有很强的自恢复能力。图 5-6 分析了规模为 n 的网络，在有 x 个节点被俘的情况下，正常网络节点通信信息可能被破解的概率，如下式所示：

$$P_{Q_becracked}=\sum_{i=q}^{m}\left(\left(1-\left(1-\frac{m}{|S|}\right)^x\right)^i * \frac{p(i)}{P_{est}}\right)$$

而 E－G 协议同等条件下被破解的概率如下式所示：

$$P_{EG_becracked}=1-\left(1-\frac{m}{|S|}\right)^x$$

其性能比较如图 5-6 所示，仿真条件为：$m=200$，任意节点对密钥建立概率 $P_{low}=0.33$。由该图可知，当被俘节点数量较少时，q－composite 模型相对于基本模型体现出更好的安全性，而随着被俘节点的增多，q 越大，其性能越差，为达到同样的 P_{low}，S 就越小；因此捕获的节点一多，就更容易恢复出 S 的内容。$q=1$ 与 E－G 协议差不多，但安全性能始

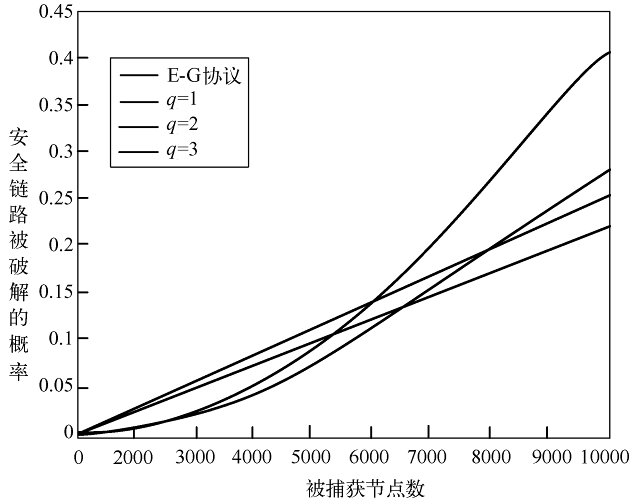


图 5-6　通信被俘比例与被俘节点之间的关系

终要好一些，这是因为 E－G 协议只用一个共享密钥来建立安全链路，而 q－composite 模式用全部共享密钥来建立安全链路。q－composite 模式以额外的计算负载为代价来提高安全性能，只适合少数节点被捕获的情况。

3. 基于地理信息或部署信息的随机密钥预分配方案

在一些特殊的应用中，节点的位置信息或部署信息可以预先大概估计并用于密钥管理，如一种针对静态 WSN 的基于地理信息的最靠近配对密钥方案 CPKS（Closest Pairwise Keys Scheme）。该方案在网络部署前，每个节点随机与最靠近自己期望位置的 c 个节点建立配对密钥。例如，对于节点 u 的邻居节点 v，部署服务器随机生成配对密钥 k_u，v，然后把（v，k_u，v）和（u，k_u，v）分别分配给 u 和 v。部署后，相邻节点通过交换节点标识符确定双方是否存在配对密钥。

CPKS 方案的优点是每个节点仅与有限个相邻节点建立配对密钥，网络规模不受限制；配对密钥与位置信息绑定，任何节点的受损不会影响其他节点的安全。其缺点是密钥连通概率的提高仅能通过分配更多的配对密钥实现，受到一定的限制。

针对上述问题，Liu 提出了使用基于地理信息的对称二元多项式随机密钥预分配方案（Location－Based Key Predistribution，LBKP）。该方案把部署目标区域划分为若干个大小一致的正方形区域。部署前，部署服务器生成与区域数量相等的对称 t 阶二元多项式，并为每一区域指定唯一的二元多项式。对于每一节点，根据其期望位置来确定其所处区域，部署服务器把与该区域相邻的上、下、左、右 4 个区域及节点所在的区域共 5 个二元多项式共享载入该节点。部署后，两个节点若共享至少一个二元多项式就可以直接建立配对密钥。该方案通过调整区域的大小来解决 CPKS 方案存在的连通概率受限的问题。与 E－G 方案和 q－composite 相比，LBK 方案的抗毁性明显提高，但缺点是计算和通信开销过大。

在基于部署知识的随机密钥预分配方案中，假定网络的部署目标区域是一个二维矩形区域且节点部署服从 Gaussian 分布。节点被划分为 $t \times n$ 个部署组，每个组 $G_{i,j}$（$i=1,\cdots,t;j=$

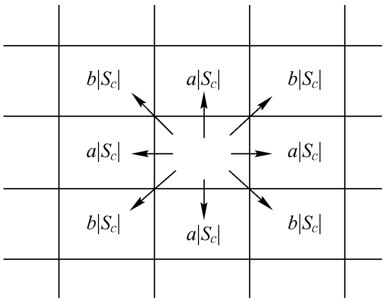


图 5-7　相邻密钥池之间的共享密钥数

$1,\cdots,n$)的部署位置组成一个栅格。密钥池（密钥数为 $|S|$）被划分成若干个子密钥池（密钥数为 $|S_c|$），每个子密钥池对应于一个部署组。若两个子密钥池水平或垂直相邻，则至少共享 $a|S_c|$ 个密钥；若两个子密钥池对角相邻，则至少共享 $b|S_c|$ 个密钥（a,b 满足以下关系：$0<a$，$b<0.25$ 且 $4a+4b=1$）。若两个子密钥池不相邻，则没有共享密钥，如图 5-7 所示。

对于组内每一节点，从对应的子密钥池随机取 m 个不同的密钥。部署后，若相邻节点存在共享密钥，则可以直接建立配对密钥。实验表明，在同等条件下，该方案提高了节点的连通概率。例如，当节点预分配的密钥数为 100 时，E-G 方案的节点连通概率仅为 0.095，而该方案能够达到 0.687。使用部署知识使得节点减少了预分配无用密钥的数量，提高了网络抗毁性。但该方案的子密钥池的划分需要慎重考虑。

尽管 Liu 和 Du 等都在密钥预分配时使用节点的位置信息以提高抗毁性，但存在攻击者容易对节点进行定位后俘获及节点因缺乏认证机制而被伪造等问题。针对上述问题，Huang 的栅格组部署方案使用限制组的节点数量、设定密钥空间被选中的阈值等方法提出了解决方案。

5.3.3　分簇传感器网络的密钥管理机制

在分簇式无线传感器网络结构中根据各个节点的功能和能量不同可以将节点分成三类：基站、簇头和成员传感器节点。分簇传感器网络的分布如图 5-8 所示。

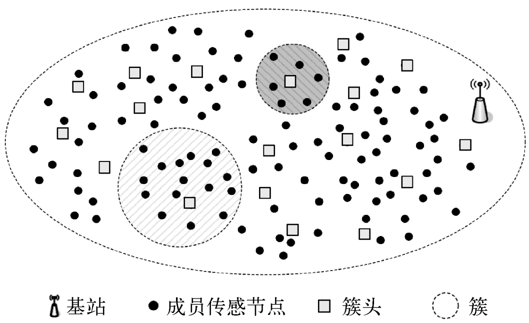


图 5-8　分簇传感器网络分布图

在网络中基站的能量和存储能力是不受限制的，它主要负责收集和处理传感器节点发送来的数据及管理整个网络。在大多数的应用中，假定基站是安全的、可以信任的，因此把基站用作密钥服务器。相对于传感器节点，簇头拥有较高的信息处理和存储能力，它负责将节点分簇、收集并处理来自节点的信息，然后将信息发送给基站。在部署网络时，传感器节点被随机地投掷在目标区域，随后节点搜寻自己无线范围内的邻近簇头自组织形成网络。

簇内通信与簇间通信关系有以下 5 种：

① 一般节点⟷一般节点；

② 一般节点⟷簇头；

③ 一般节点⟷基站；

④ 簇头⟷簇头；

⑤ 簇头⟷基站。

到底需要哪些密钥取决于无线传感器网络中数据的加密策略和数据融合的需要。在数据的传输过程中如果中间节点对经过传送的数据包内容不做任何处理，仅仅是起到一个疏导的作用，则可以采用端到端的加密策略——仅源节点与目的节点共享密钥就行了。相反，如果中间节点对经过它传送的数据要进行处理、加工，特别是无线传感器网络，为节能、提高数据准确性和数据收集效率，要采取数据融合措施，这时就要采取点到点的加密策略，数据传输途径中，每相邻节点间都应共享密钥。在大多数情况下，簇头应是该簇的数据转发中心和数据融合中心。因此，我们可以把分簇的无线传感器网络的密钥管理技术归纳如下。

簇内：只需要每个节点与该簇的簇头共享密钥。如果中间需要经过多路径，则采用簇内节点 – 簇头的端 – 端传送方式，不考虑簇内节点之间的密钥共享。

簇间：有两种情况，如果每个簇送到基站的位置是独立的，不需要经过其他簇头进行数据融合，则只要每个簇头与基站共享密钥就行了；如果各簇的数据之间还要进行数据融合，则需要每个簇头与邻近的簇头共享密钥。

因此研究分簇的传感器网络时，对于簇内密钥分配与簇间密钥分配应采用不同的策略。另外，传感器网络簇头选取方式不同，也导致密钥分配方式有很大不同。

1. LEAP 密钥管理方案

2003 年 Zhu 等人提出的 LEAP（Localized Encryption and Authentication Protocol）是一个既能支持网内处理，又具有较好抗捕获性的密钥管理协议，这种协议支持四类密钥的生成和管理，提供了较好的低能耗的密钥建立和更新方案，同时还提供了基于单向密钥链的网内节点认证方案，并在不丢失网内处理功能和被动参与的情况下支持源认证操作。

Zhu 等人认为应该在网络节点中设立多种密钥以适应不同的需要，因此在 LEAP 中建立了四种类型的密钥：个体密钥、对密钥、簇密钥、组密钥。每种密钥都有不同的作用，各个密钥建立过程的具体步骤如下。

1）个体密钥

个体密钥为节点与基站所共享的密钥，由节点在部署钱通过预分配的主密钥和伪随机函数 f 来生成，用于节点向基站发送秘密信息。节点 u 的个体密钥产生公式如下：

$$K_u^m = f_{K_m}(u)$$

2）对密钥

对密钥是相邻节点单独共享的密钥，用于节点间单独交换秘密信息，是通过交换其标识符及使用预分配的主密钥和单向散列函数计算得到的，具体产生步骤如下。

① 密钥预分配。管理节点产生一个初始化密钥 K_I，每个节点预存 K_I，并按下式计算出节点自身的主密钥：

$$K_u = f_{K_I}(u)$$

② 邻居发现。部署后，节点广播自己的标识符 ID，邻居节点接收到信息后回复源节点，

格式如下：

$$u \rightarrow * : u$$

$$v \rightarrow u : v, \mathrm{MAC}(K_v, u \mid v)$$

③ 对密钥建立。节点收到邻居节点的回复后就可以计算对密钥，按下式计算；

$$\begin{cases} K_{uv} = f_{K_v}(u), u < v \\ K_{uv} = f_{K_u}(u), u \geqslant v \end{cases}$$

④ 密钥撤销。对密钥建立周期过后，每个节点 u 撤销 K_I 及所有 K_v。

3）簇密钥

簇密钥为同一簇内相邻节点所共享，由簇头产生一个随机密钥作为簇密钥，然后使用与邻居节点的对密钥逐一把簇密钥加密后发送给邻居节点，邻居节点把簇密钥解密后保存下来。

4）组密钥

组密钥为基站与所有节点共享的通信密钥，基站首先利用簇密钥将组密钥加密，并将其广播给自己的子节点，子节点获取最新的组密钥后用与其下一级子节点共享的簇密钥加密组密钥后广播给其子节点。以此类推，直到所有节点都获取最新的组密钥为止。

在 LEAP 管理机制中，任何节点的受损都不会影响其他节点的安全。但也存在一定的缺点，节点在部署后，在一个特定的时间内必须保留全网通用的主密钥。主密钥一旦被暴露，则整个网络的安全都受到威胁。此外，在对密钥生成阶段因为只有单向认证，因此还存在 hello 攻击，即当攻击者 t 假冒除 v 外的网络中的任何节点向节点 v 广播协商请求时，按照协议节点 v 将生成对所有节点的对密钥。

2. 基于 EBS（Exclusion Basis Systems）的动态密钥管理方案

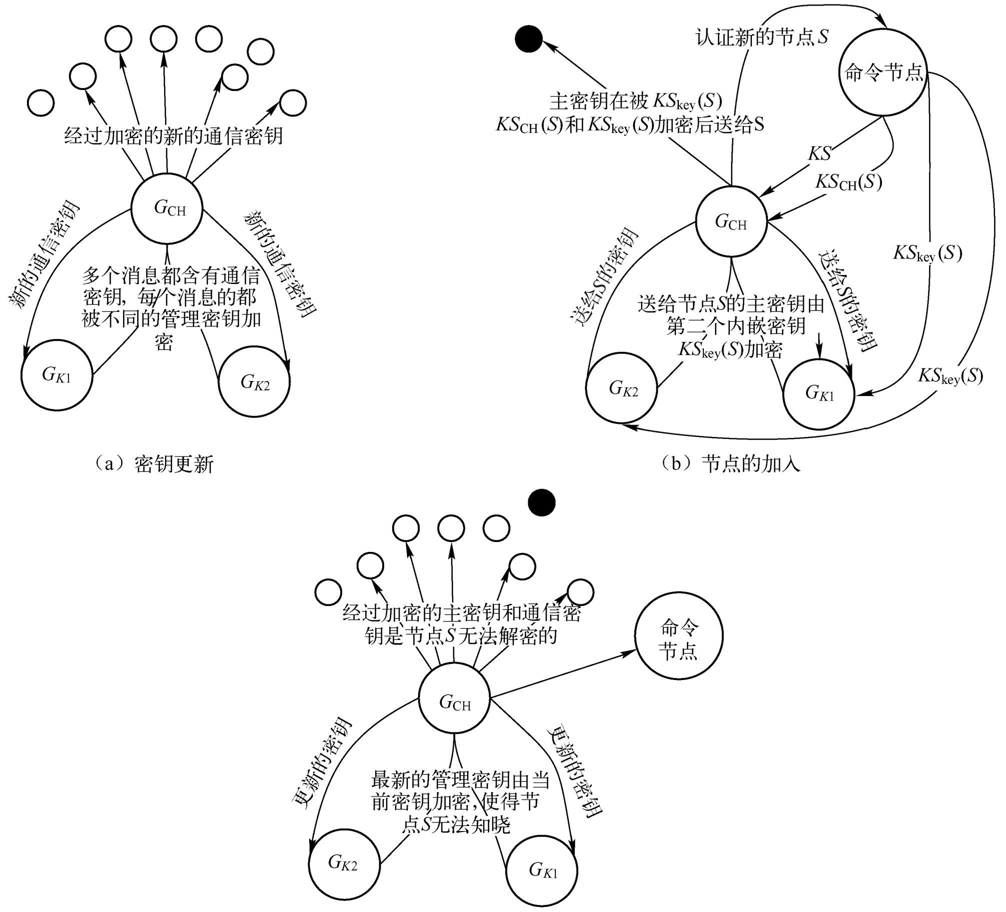
EBS 由 Eltoweissy 提出，主要用于密钥动态管理。EBS 为一个三元组（n,k,m）表示的集合 Γ，其中，n 为组的用户数，k 为节点存储的密钥数，m 为密钥更新的信息数。对于任一整数（用户）$t \in [1,n]$，具有以下属性：

① t 最多出现在 Γ 的 k 个子集（密钥）里，表示任一用户最多拥有 k 个密钥；

② 有 m 个子集（密钥），$A_1, A_2, \cdots, A_m$，满足 $\bigcup_{i=1}^{m} A_i = [1,n] - \{t\}$，表示使用 m 个与 t 无关的密钥更新信息可撤回用户 t。

Younis 在层次式 WSN 里提出了基于位置信息的 EBS 动态密钥管理方案 SHELL（Scalable，Hierarchical，Efficient，Location－aware，and Light－weight）。在 SHELL 方案里，普通节点按照其地理位置被划分为若干簇，由簇头或网关节点（Gateway）来控制。网关节点有可能被命令节点指定为其他簇的密钥生成网关节点（Key Generating Gateway）。它并不存储和生成自己簇里各节点的管理密钥。根据簇数和节点的存储容量，簇 C_i 的网关节点 $G_{\mathrm{CH}}[i]$ 使用正则矩阵法生成所在簇的（n,k,m）——EBS 矩阵，并把矩阵的相关部分内容分别发送给该簇的密钥生成网关节点 $G_{K1}[i]$ 和 $G_{K2}[i]$ 等。密钥生成网关节点根据 EBS 矩阵的内容生成相应的管理密钥，并通过网关节点 $G_{\mathrm{CH}}[i]$ 广播给簇内各节点。为了避免串谋攻击，相邻节点管理密钥的汉明距离（Hamming distance）设计为最小。

SHELL 定期更新密钥。当需要更新密钥时，由簇头首先把最新的通信密钥发送给密钥生成网关节点，然后由密钥生成网关节点生成新的管理密钥，再通过簇头发送给簇内各节点，如图 5-9（a）所示。



（a）密钥更新　　（b）节点的加入

（c）受损节点的撤回

图 5-9　密钥更新、节点的加入与受损节点的撤回

当新的节点加入时，首先根据其地理位置确认加入所在簇，并通过命令节点认证其身份，然后由簇头与密钥生成网关节点协调启动管理密钥生成进程，如图 5-9（b）所示。当要撤回受损的节点时，若是簇头受损，则可以采取指定新的簇头或把簇内节点重新分配到其他正常的簇内等方法；若是普通节点受损，簇头把受损节点信息通知密钥生成网关节点，然后由密钥生成网关节点利用 EBS 的性质生成新的管理密钥，并通过簇头广播发送给簇内节点，受损节点由于无法解密广播数据包而无法获取新的管理密钥，如图 5-9（c）所示。

与随机密钥分配方案相比，SHELL 明显增强了抗串谋攻击的能力。例如，当 $k=4$，$n=200$ 时，若要发起串谋攻击，则在 SHELL 里需要使 11 个节点受损，而在随机密钥分配方案时仅需要 3 个节点受损。但在 SHELL 里由密钥生成网关节点存储相应簇的节点密钥，这意味着，密钥生成网关节点受损数量越多，网络机密信息暴露的可能性就越大。针对 SHELL 的缺点，Eltoweissy 提出了 LOCK（Localized Combinatorial Keying）方案。该方案使用两层

EBS 管理密钥对基站、簇头和普通节点的密钥分配、更新、撤回进行管理，使得簇头的受损不会暴露更多的机密信息。

5.3.4　基于 PUF 的 DTMSN 密钥管理机制

在 4.3.4 节介绍 MAP 机制的基础上，这里提出一种基于 PUF 的延迟容忍移动传感器网络密钥管理机制（Key Management scheme based on PUF，KMP），该机制采用分层的网络结构，即除了基站外，传感器网络还包括两类节点：摆渡节点和终端节点，其中摆渡节点主要用来收集终端节点的信息并上传给基站，以降低终端节点的存储量。分层的结构有助于提高节点管理的效率和降低密钥管理的难度。分析表明，该机制能够有效抵抗节点克隆等多种类型的网络攻击。

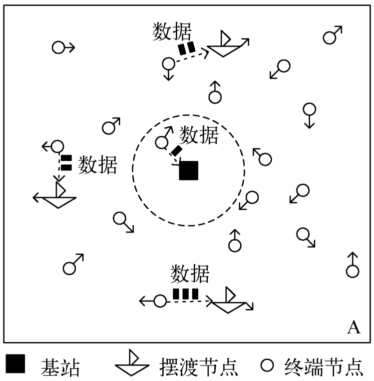


图 5-10　KMP 网络模型

1. 网络模型及假设

KMP 网络模型如图 5-10 所示，在监测区域 A 中部署了多个 DTMSN 节点（箭头代表运动方向）。节点类型共有三种，其中基站主要用来完成密钥的管理，以及感知信息的收集；另外两类节点分别是摆渡节点（Ferry）和终端节点（Endpoint），统称为传感器节点。摆渡节点具有更多的能量和存储空间，主要用来辅助终端节点将信息转发到目标节点；终端节点能量及资源都非常有限，主要用于完成信息的采集。网络中基站静止不动，摆渡节点和终端节点可以运动；摆渡节点和终端节点芯片上都具有 PUF 单元，且终端节点的数量要远大于摆渡节点。

为了提高数据传输成功率，以往的 DTMSN 网络中，节点间通常彼此转发消息，这种转发常常导致节点的有限存储空间被迅速填满，反而降低了消息的传输成功率，如 Flood 传输机制。在 KMP 网络模型中，终端节点只与基站和摆渡节点通信，终端节点可以通过摆渡节点将数据传输给基站；摆渡节点不但可以在终端和基站之间进行消息转发，也可以在终端节点之间进行消息转发。除了以上对模型的描述外，KMP 还有类似于 MAP 的如下假设。

① 网络中每个节点的芯片上都内嵌 PUF 单元，该 PUF 单元与芯片密不可分，任何分离 PUF 与芯片的操作或破解芯片的操作都将导致 PUF 单元的破坏。

② 基站在网络中被认为是可信的，并且不受资源和能量的限制；摆渡节点的能量和存储资源可以认为是无限的；而终端节点的资源和能量都有限。

③ 假设攻击者能力足够强大，不但能够侦听网络数据、向网络注入报文、实施重放攻击等操作，而且对于捕获的传感器节点（摆渡节点和终端节点）能够实施破解，获取存储在芯片中的信息。但基站是安全的，不能被破解。

2. 密钥分类

根据网络通信过程中不同的安全需求，参考 LEAP 安全机制，将网络中的密钥分为三类，具体如下。

私有密钥：每个摆渡节点、终端节点与基站之间都共享一个私有密钥，这个密钥主要用来实现摆渡节点、终端节点与基站之间的安全通信，因此摆渡节点转发的终端节点与基站间的消息对于摆渡节点来说是机密的。

全局密钥：网络中所有节点共享的一个密钥，主要用于基站进行全网广播消息的加密。

挑战密钥：摆渡节点与任意一个终端节点都共享一个挑战密钥，利用挑战密钥终端节点间可以与摆渡节点实现加密通信，也可以在摆渡节点转发的基础上实现终端节点间的加密通信。同时，挑战密钥还可以作为认证过程中的 PUF 的挑战，用来生成认证数据。

从这三类密钥可以看出基站的级别最高，摆渡节点次之而终端节点最低，因此终端节点间的通信数据可以被摆渡节点看到，而终端节点与基站的通信则对摆渡节点不可见。在介绍 KMP 之前，首先列出需要用到的表达式，参见表 5-1。

表 5-1　符号及表达式说明

GK	全局密钥
IK_k^l	经过 l 次更新后的传感器节点 k 的私有密钥
CK_{ij}^l	经过 l 次更新后传感器节点 i 和 j 之间的挑战密钥
$\parallel$	连接操作符
$\rightarrow$	单播
$\Rightarrow$	广播
$\oplus$	异或操作
$H(s)$	对消息 s 的 Hash 操作
$P_i(C)$	传感器节点 i 相对于挑战 C 的 PUF 响应
$\mathrm{MAC}(k,s)$	利用密钥 k 对消息 s 进行的消息认证码操作

3. 初始化密钥分配

假设 DTMSN 网络中有 n 个终端节点和 m 个摆渡节点，其中 $m \ll n$，传感器节点 k 的身份标识为 ID_k，k 为整数；基站节点的身份标识设为 SID。网络部署前，基站为每个传感器节点生成和预分配一系列参数，同时也将这些参数存储在自身的存储单元中，各参数具体如下。

初始挑战密钥的建立：首先基站为终端节点 i 和摆渡节点 j 随机生成一个共享的初始挑战 $\mathrm{CK}_{ij}^0 \in \mathfrak{R}^n$，同时利用节点各自的 PUF 单元获得初始密钥的挑战响应 $P_i(\mathrm{CK}_{ij}^0)$ 和 $P_j(\mathrm{CK}_{ij}^0)$。接下来，基站选择单向 Hash 函数 H（如 MD5），并为终端节点 i 和摆渡节点 j 计算 $H(P_j(\mathrm{CK}_{ij}^0))$ 和 $H(P_i(\mathrm{CK}_{ij}^0))$。最后基站将函数 H 和三元组 $<\mathrm{ID}_i, \mathrm{CK}_{ij}^0, H(P_i(\mathrm{CK}_{ij}^0))>$ 下载到摆渡节点 j 中，将函数 H 和三元组 $<\mathrm{ID}_j, \mathrm{CK}_{ij}^0, H(P_j(\mathrm{CK}_{ij}^0))>$ 下载到终端节点 i 中。因此，最后网络中每个终端节点将存储 m 个摆渡节点的三元组信息；而每个摆渡节点将存储 n 个终端节点的三元组信息，同时为了使摆渡节点之间也能进行安全通信，摆渡节点也需要存储其他 $m-1$ 个摆渡节点的三元组信息，这样摆渡节点共存储 $n+m-1$ 个三元组信息。

初始私有密钥的建立：任意一传感器节点 k 的初始私有密钥建立的方法与挑战密钥类似。首先，基站选择一个随机数 $\mathrm{IK}_k^0 \in \mathfrak{R}^n$ 作为 k 的初始私有密钥，并且以此为挑战从节点 k 及自身的 PUF 单元获取响应 $H(P_{\mathrm{Sink}}(\mathrm{IK}_k^0))$ 和 $H(P_k(\mathrm{IK}_k^0))$，并将三元组信息 $<\mathrm{SID}, \mathrm{IK}_k^0, H(P_{\mathrm{Sink}}(\mathrm{IK}_k^0))>$ 下载到节点 k 中。这样，每个传感器节点（摆渡节点和终端节点）都要多存

储一个关于基站的三元组信息。

全局密钥的建立：全局密钥 GK 主要用于基站进行机密消息广播，是被网络中所有节点都共享的一个密钥。GK 由基站随机生成并下载到网络中所有节点的，其建立方法非常简单。

4. 节点间的相互认证

节点间的相互认证主要在终端节点与摆渡节点、终端节点与基站节点、摆渡节点与摆渡节点、摆渡节点与基站节点之间发生。具体的认证过程与 4.3.4 节中介绍的认证过程相同，不再赘述。任意两个节点之间的第 $k+1$ 次相互认证过程如图 5-11 所示，$CK_{ij}^{k+1}=H(CK_{ij}^{k})$。同样，每次进行数据传输之前节点间都要进行认证，而且认证过后都需要进行参数的更新，具体的更新方法与 4.3.4 节中介绍的相同。

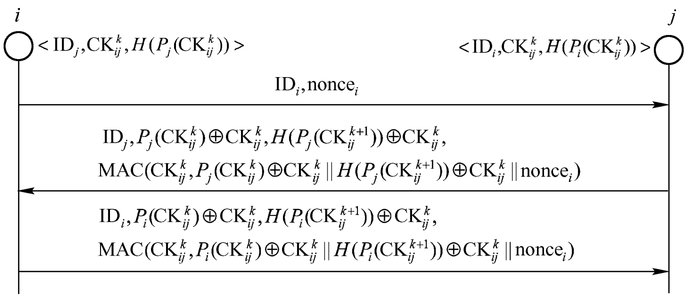


图 5-11　节点 i 和 j 之间的 $k+1$ 次认证过程

由于终端节点之间不直接进行数据传输，因此彼此之间不进行认证操作。当终端节点与摆渡节点与基站直接通信时，此时认证中的挑战使用的是 IK_i^k 而不是 CK_{i*}^k。

5. 密钥更新

由于网络采用基于 PUF 的安全密钥管理机制，为了提高网络安全性，防止 PUF 信息被重复利用，需要对网络中的密钥进行及时更新。各密钥更新的方法如下。

挑战密钥更新：挑战密钥更新的方法实际上与挑战值的更新方式相同，如在第 k 次成功完成节点间的相互认证后，节点 i 和节点 j 间的挑战密钥由原来的 CK_{ij}^{k-1} 变为 CK_{ij}^{k}，其中 $CK_{ij}^{k}=H(CK_{ij}^{k-1})$，密钥的更新有利于防止攻击者根据认证信息实施克隆攻击。

私有密钥更新：通常情况下，私有密钥主要用来实现传感器节点发送给基站的消息的加密。当传感器节点与基站直接通信时，它们之间也需要有认证过程来验证对方身份的合法性。具体的私有密钥更新过程与挑战密钥一样，即 $IK_i^k=H(IK_i^{k-1})$。

全局密钥更新：当网络处于危险环境中时，如某些节点被攻击者破解，此时网络的全局密钥需要进行更新，即便是网络环境良好，出于安全考虑也应该周期性地更新全局密钥。基于网络节点的移动性，对于一个新的全局密钥 GK_{new}（由基站随机生成），在网络中分配的方法非常简单：一方面，基站可以利用私有密钥对全局密钥进行加密，然后传输给与其发生连接的传感器节点；另一方面，接收到新的全局密钥的摆渡节点可以利用挑战密钥对全局密钥进行加密，并将其传输给与其发生连接的终端节点。

从以上的介绍中可以看出，KMP 中密钥的更新方法非常简单，主要原因是它充分利用了网络节点的移动性和分层的网络结构。

6. 密钥重置及网络扩展

一般而言，在网络部署以前，基站会对每个节点采集足够多的挑战－响应对。这些挑战－响应对可以用于后期的网络维护。在一些极端的网络环境中，由于客观环境或攻击者使得节点间的挑战密钥更新不同步时，往往造成下一次节点间的认证无法正确进行，这时，基站可以采用初始化时的相关操作将新的挑战密钥下载到节点中。相反，如果传感器节点的私有密钥没有与基站同步更新，此时基站可以利用存储的传感器节点的信息自动进行调整，使其与节点的私有密钥相同；如果私有密钥不得不重新设置，则基站可以采用网络初始化时的相关操作将新的私有密钥下载到节点中。

当网络中有新的节点加入时，有关新节点的各类密钥设置的具体方法为：假设网络中有新的终端节点 i 加入，首先，基站将当前网络全局密钥 GK、基站随机生成的初始私有密钥 IK_i^0 及 H 函数下载到新的节点中；其次，基站根据在网络部署前已经存储的有关摆渡节点 j 的一个挑战响应对$(C, P_j(C))$，选择 C 作为新节点 i 与摆渡节点 j 之间的挑战密钥，并在获得新节点的响应 $P_i(C)$后将三元组$<\mathrm{ID}_j, C, H(P_j(C))>$下载到新节点中，至此新加入节点 i 的密钥信息全部建立。而基站会在与摆渡节点建立连接时将有关新节点 i 的相关信息传输给摆渡节点。同样，当网络中有新的摆渡节点加入时，采用的密钥分配的方法类似，只不过此时需要将有关新的摆渡节点的三元组信息传输给所有的传感器节点，具体传输方法有很多，不是本书研究的重点，这里也就不再赘述。

7. KMP 的安全性分析

KMP 的安全性同样是基于 PUF 单元的，下面给出该机制能够有效抵御的主要典型网络攻击。

① 节点克隆攻击：DTMSN 网络中，传感器节点很容易被攻击者发现并进行破解，而网络维护者却很难发现。而当攻击者对节点进行复制后，由于 PUF 无法复制，则不能通过其他节点的认证，也就无法获取节点传输的消息。可见 KMP 机制对于节点克隆攻击具有很好的抵御能力，能够有效降低克隆攻击带来的安全威胁。

② 节点伪装攻击：在 KMP 机制中，由于所有节点通信前都需要进行身份认证，因此任何伪装其他节点的企图都将被识别。因此，KMP 能够有效抵御节点伪装攻击。

③ 窃听攻击：由于网络中的所有通信消息都由密钥进行加密后进行通信，因此攻击者无法通过窃听获取网络信息。

④ 破解攻击：在 KMP 中，由于传感器节点的私有密钥和挑战密钥各不相同，因此即便节点被攻击者捕获并破解，其内部存储的全局密钥、挑战密钥及私有密钥都泄露，此时攻击者唯一能够捕获的就是网络全局广播的消息，而因为全局密钥周期性更新，因此他能够获取的消息通常较少。而且攻击者无法利用挑战密钥和私有密钥对其他节点实施攻击，因为这些密钥只是针对被捕获节点的，对其他节点不构成威胁。因此，KMP 能够有效抵御破解攻击。

从上面的分析中可以看出，KMP 由于采用了多种密钥，从而有效地减少了节点被破解后对网络产生的威胁，而且由于 PUF 的不可复制性，也使得传感器节点能够抵抗攻击者的克隆攻击，这对于其他密钥管理机制而言是不具备的，因此从安全性上来看，KMP 能够有

效地保证延迟容忍移动传感器网络的安全性。

5.4　基于物理层信道特征的密钥生成技术

5.4.1　物理层安全

物理层安全的基本思想：利用通信信道噪声的随机性及应用通信干扰技术、编码技术等确保恶意窃听用户无法获取发送消息的任何有用信息。换言之，物理层安全就是利用噪声和信道特征固有的随机性，限制非认证接收者获取有用信息。

1. 物理层安全分类

基于信息论的物理层安全主要包含两个分支：由 Wyner 所倡导的无密钥安全和由 Shannon（香农）和 Maurer 所倡导的基于密钥的安全机制（见图 5-12），这里主要关注后者。物理层安全利用噪声信道内在的不确定性，而非复杂的数学计算来实现信息传输的机密性。传统密钥加密的手段从网络协议的角度来提供数据机密性服务，而物理层安全则是从信息论和信号处理的角度来寻求安全服务。

基于密钥安全机制的无线物理层安全问题按照对于理论和实践的侧重点的不同，研究大体上可分为两大类：一是从信息论的角度对无线网络的安全信道容量进行理论分析，得出不同信道类型条件下的信道容量的理论上（下）限值；二是从信号处理和优化的角度，研究系统的安全通信速率逼近其安全容量的实现途径。

理论研究的一个重要概念是安全容量。安全容量被定义为：保密信息被目标接收端可靠接收，并且非法接收端无法获取任何有用信息的最大可达通信速率。在非退化离散信道中，它等价于通信信道容量与窃听信道容量之间的差值。

密钥生成速率的界线：相互观测随机过程用于密钥生成，如信道冲击响应，通过公开的广播信道，在窃听者 Eve（简记为 E）存在的情况下，Alice（简记为 A）和 Bob（简记为 B）能够得到的密钥速率。

上限：$S(A,B|E)\leqslant\min[I(y_A;y_B),I(y_A;y_B|y_E)]$；

下限：$S(A,B|E)\geqslant\max[I(y_A;y_B)-I(y_A;y_E),I(y_A;y_B)-I(y_B;y_E)]$

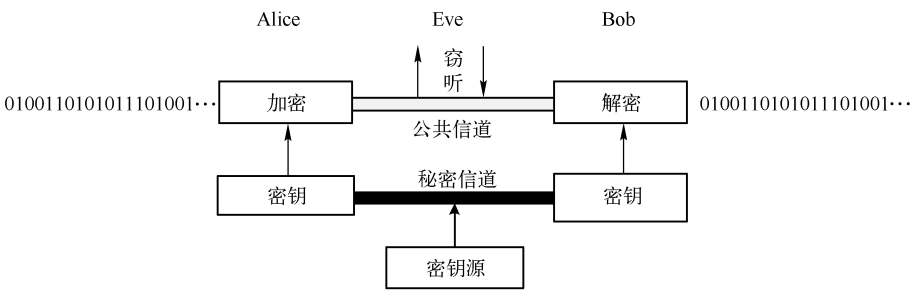


图 5-12　香农的保密通信系统模型

当窃听者 E 与合法通信者 A 和 B 间的信道没有相互信息时，上下限间将变得很紧凑，这种情况往往通过使窃听者远离合法通信者而实现。无线信道安全容量的分析是物理层安全

问题研究的基础，它为物理层安全的实践研究提供了理论依据。

物理层安全的实践研究细分为三个子方向：第一，安全编码，优化设计适应于信道环境的信道编码方式，实现合法用户能够成功译码而窃听者不能的目的；第二，加扰方法，人工发送特定干扰信号，降低非法用户处信号的信噪比来破坏其译码成功率，而对于合法用户免疫；第三，信道特征加强安全服务，包括利用无线信道特征的密钥构建技术和利用无线信道特征的认证技术，合法用户通过共享信道特征实现密钥分发及通过对信道特征参数的评估来鉴别前后通信者的身份，后者通常用于对恶意攻击行为的检测。

2. 物理层安全与传统密码安全机制的区别

物理层安全与传统的密码机制安全无论从实现方式还是安全衡量标准方面都存在明显区别，具体表现如下。

首先，作用于 OSI 模型的协议范畴不同。传统的密码机制安全作用于网络协议栈的高层，而基于信息论的物理层安全处于网络协议栈的底层。

其次，作用机理、实现代价不同。传统的密码机制安全属于计算安全手段，通常依赖公钥密码系统实现用户身份识别认证、密钥分配及对称加密以确保数据的机密性，并以算法的高计算复杂度为代价。其中，传统的密钥分配机制建立在计算模型基础上，需要可信第三方参与，协议及系统架构都较为复杂。并且假定单向函数无法逆转，而这个假定至今尚未得到严格的数学证明。随着计算能力的提升，如量子计算机技术的发展，使得基于计算复杂度的安全模式受到威胁。依信息论观点来看，基于公钥的加密系统无法提供信息论角度上的信息安全。因其假定所有信道均是无噪信道，所以通信系统的安全速率为零。相比之下，物理层安全无须复杂度很高的算法，通过在物理层对信息进行处理，可从根本上确保信息的安全性。

再次，安全评估可行性不同。传统的密码机制安全难以制定一个精确的标准来衡量不同加密算法的安全性能，通常只是简单利用系统遭受攻击次数的多少来衡量。相对而言，基于信息论安全模型的物理层安全易于从理论上来量化分析信息泄露的程度，对于系统的安全防护级别的评估更加准确、简单易行。

因此，针对物理层安全问题的研究是无线网络信息安全发展的必然趋势，是进一步提高无线网络信息安全性的根本途径。已有研究表明利用无线信道特征构建动态密钥可用于现存加密机制。

3. 无线信道的传输特性

无线信道是无线通信中发送机和接收机间通信链路的一种抽象表述。信息在信道中依托电磁波进行传播，传播特性是无线通信系统首要研究的问题。电磁波在空间中的传播路径，可粗略地划分为不受阻挡的直视路径和存在阻挡物体的非直视路径。电磁波在直视路径上表现为直线传播形式，而在非直视路径表现为反射、绕射和散射传播方式。无线传播具体的传播形式取决于无线网络所处的环境且随时间而改变，具有复杂性和时变性。信道对信号产生的衰落影响可按时域、空域、频域三个角度进行描述。

1）时域特性

时域内，无线传输衰落按照功率降低的速率可分为慢衰落和快衰落。慢衰落表征信号幅

度在较长时间、较大范围内的变化情况，故又称长期衰落或大尺度衰落；快衰落表征信号幅度的瞬时、小范围内的变化情况，故又称短期衰落或小尺度衰落。形成大尺度衰落的原因主要是路径损耗和阴影效应，而小尺度衰落的原因是多径效应的存在。无线传输衰落的时域表现如图 5-13 所示。

大尺度衰落表征接收信号功率的均值与传输距离的量值关系。其中，路径损耗指电磁波在空间传播时所产生的损耗，主要是功率的辐射扩散所致，表现为较大范围内接收信号功率均值的变化。阴影效应指由于电磁波受到大型建筑物及其他物体的阻挡而形成电磁场阴影，随着移动台位置改变而产生的接收点场强中值的幅值变化现象。小尺度效应是指短距离（几个波长以内）或短时间（秒量级）内接收信号经历的剧烈变化。

大尺度衰落效应的影响可以被利用以实现身份认证，准确地说，是利用大尺度衰落效应实现对通信信道的区分认证。基于信道特征参数构建密钥基于小尺度衰落的影响，应该避免大尺度衰落的影响，防止敌手利用与合法通信用户位置有关的大尺度衰落的影响进行密钥预测。

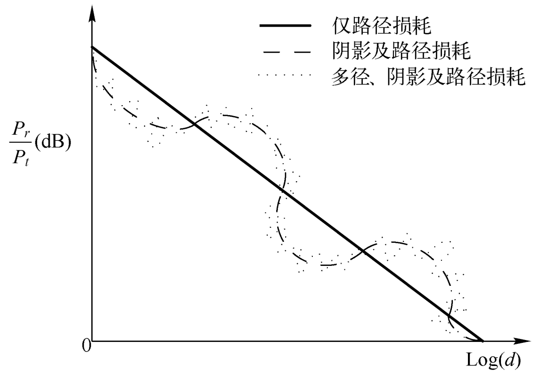


图 5-13　无线环境中衰落效应时域表现图

2）空域特性

空域内，无线传输衰落根据是否存在直视路径可分为瑞利衰落和莱斯衰落。在瑞利衰落中，传输路径中不存在占支配地位的直视路径，各个路径间彼此相互独立。换言之，瑞利衰落中不存在直射波信号，接收信号仅由反射波、折射波、散射波等信号所组成，此时接收信号的包络幅值服从瑞利分布；在莱斯衰落中，在多个传输路径中存在一条直视路径，其信号属性占支配地位，接收端表现为可接收到较强的直射波，此时接收信号的包络幅值服从莱斯分布。

当主路径接收的信号强度表现远远大于其他路径的信号时，多径效应应当忽略，此时接收信号的幅值将近似表现为正态分布。

3）频域特性

频域内，无线传输衰落按照衰落与频率的对应关系分为平坦衰落和频率选择性衰落。若信道频带比信号频谱范围宽，多重路径信道中的传送信号的频谱大致保持不变，则信号传输表现为平坦衰落；若多路信号的相对时延与一个符号的时间相比不可忽略，当多路信号叠加时，时延扩展导致不同时间的符号重叠在一起，造成符号间的干扰，则信号传输表现为频率选择性衰落。平坦衰落和频率选择性衰落分别对应无线系统中的窄带信道和宽带信道。

频率选择性衰落不仅导致接收信号的幅值产生随机变化，信号的波形也可能产生失真。因此，频率选择性衰落对传输质量的影响更大。频率选择性衰落的发生取决于信道和信号两个方面。当信道带宽远大于信号带宽时，且频带内具有恒定的增益和线性相位响应，则接收信号只发生平坦衰落。此时，信道的多径结构可保证接收端保持发射信号的频谱特性。但由于多径引起的信道增益扰动，接收信号的强度表现往往随时间的变化而变化。当上述条件不成立时，接收信号表现为频率选择性衰落。

此外，接收机与发射机之间的相对移动引入了多普勒频移。由于通信终端本身或周围物体的移动，接收信号频率发生变化的现象称为多普勒效应，多普勒效应所引起的附加频移称为多普勒频移，其与通信终端的运动速度、运动方向及接收机处多径电磁波的入射角度有关。

5.4.2　基于信道特征的密钥生成

1. 信道特征构建密钥可行性分析

如前面章节所述，由于多径传播的原因，无线环境中所接收到的信号是无限个原始传输信号衰弱、时延、相移后版本的总和，各路径间相互干扰。同时，环境变化和移动终端的相对移动都可能造成路径的改变，这些因素的共同作用将导致接收信号幅度和相位的随机变化。空谱指信号在传播空间上的分布，不同用户的空谱不同，其差异的大小决定了物理层安全性的高低。空谱刻画了以下 4 个属性，成为基于无线物理层信道特征生成和提取密钥的基础。

随机性：源信号从发送机经过复杂的多径衰落过程到达接收机，多路径的信号复合使得到达接收机的信号是传输信号的一个随机失真，具有随机源特性。由于多径的传输，接收信号是有限个发送信号因衰落、时延、相位移导致的各类复本的总和叠加，取决于各路电磁波的相位，这种叠加可能是破坏性的。接收信号取决于各路电磁信号传播时延的不同和幅度的相互关系。信道特征的随机性是信道本身所固有的，而非因采用特定机制生成的伪随机，因此对于非预定的接收者而言，是不可能通过计算来破解的。

快速时变性：无线信道状态随时间快速变化，使得信道特征在时间间隔大于信道相干时间的条件下，彼此是独立的。快速时变性有助于实现“一次一密”，同时增加敌手的侦听和破译难度。

快速空变性：依据无线电理论知相距 1/2 个波长以上的两个天线经历的衰落是不相干的结论，对处于信号若干个波长以外位置的窃听者而言，窃取有用信道信息是不可能的。由于无线信道的空间不相关性，两个位于不同位置的接收者，对于同一发送者所观测的信道特征是不同的。例如，对于频率为 2.4GHz 的无线电而言，当两个接收机相距 $\lambda/2=6.25\text{cm}$ 的距离时，二者所感知的信道特征是不同的。

短时互易性：在时分双工（Time – Division Duplexing，TDD）系统中，前向链路和反向链路采用相同的频带传输信息，在较短时间内可以认为前向链路和反向链路经历的信道特征是相似的，该性质称为信道互易性（Channel Reciprocity）。在相干时间内，信道对于同时同频通信的无线链路两端的收发器产生的衰落理论上是一致的，满足短时互易性。无线信道的互易性使得合法用户通过无线介质传输加密信息而无须交换加密密钥。互易性使得通信双方在密钥生成过程的同时实现了密钥分发。互易性是基于这样的事实，发送的电磁波在双向经历着相同的物理扰动（如反射、折射、散射等），因此，上下行链路工作在相同的频带上，

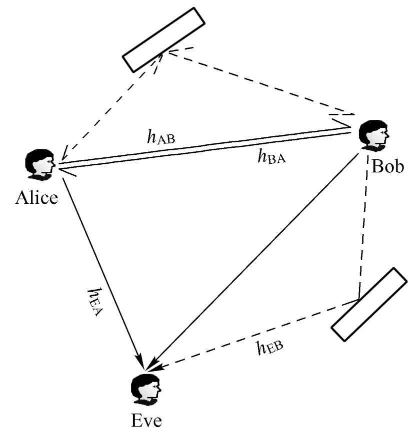


图 5-14　窃听者 Eve 存在情况下的信道互易性模型

与方向无关，信号经历相同信道响应。互易性的电磁波基础从 1896 年以来就被研究和扩展。信道满足互易性意味着一对通信者无须传输信道信息，它们可以评估和探测其间的信道特征。

因此，基于无线信道特征构建密钥是利用无线通信的本质属性，属于内生安全，能够有效解决密钥分发问题。为更直观地说明物理层密钥生成方法，有必要详细阐述其重要基础：信道互易性模型（见图 5-14），Alice（简记为 A）和 Bob（简记为 B）为处在多径衰落环境下采用 TDD 模式合法的通信节点，Eve（简记为 E）为不能产生强干扰的被动窃听者，h_{XY}代表 Y 所感知到的 X 与 Y 间的信道响应（其中 $X,Y \in S(\mathrm{A},\mathrm{B},\mathrm{E})$）。为了评估随机信道参数，Alice 和 Bob 必须交替地给对方发送探测信号 $s(t)$。Alice 和 Bob 作为接收机时接收到的信号可以表示为：

$$r_{BA}(t) = s(t) * h_{BA}(t) + n_{BA}(t) \tag{1}$$

$$r_{AB}(t) = s(t) * h_{AB}(t) + n_{AB}(t) \tag{2}$$

Eve 作为窃听者接收到的 Alice、Bob 所发送的信号分别为：

$$r_{AE}(t) = s(t) * h_{AE}(t) + n_{AE}(t) \tag{3}$$

$$r_{BE}(t) = s(t) * h_{BE}(t) + n_{BE}(t) \tag{4}$$

以上等式中，$s(t)$代表训练信号；$n(t)$为各接收机处彼此不相关的随机噪声；利用收到的信号，Alice 和 Bob 分别对信道参数进行评估。由互易性原理知，在相干时间内信道参数变化量很小，h_{AB}与 h_{BA}有极强的相似性，可假定二者相等。当 Eve 距离任一合法通信方大于 1/2 个波长时，其所感知的信道状态与该合法用户有极低的相关性，即 $h_{AE}(t)$、$h_{BE}(t)$与 $h_{AB}(t)$、$h_{BA}(t)$在幅度值变化方面是不相关的。由于 $h_{AE}(t) \neq h_{BE}(t) \neq h_{AB}(t)(h_{BA}(t))$，$n_{AE}(t) \neq n_{BE}(t) \neq n_{AB}(t) \neq n_{BA}(t)$，$r_{AE}(t)$、$r_{BE}(t)$与 $r_{BA}(t)$、$r_{AB}(t)$互不相关，Eve 不可能估计出与 Alice、Bob 完全相同的信道特征。合法节点通过量化及密钥一致性协商可以将信道参数评估值转化为相同的比特串，再利用隐私加强技术，也就是筛选一部分作为密钥，去除在协商过程中可能造成泄露的成分。必须指出，信道参数是时变的，基于信道特征参数提取出来的密钥序列也具有定时更新的动态性，这使得窃听者破解变得更加困难，保证了通信的安全需求。

无线物理层密钥研究利用信道特征的随机性（randomness）、互易性（Reciprocity）、时变性（Temporal Variations）和空变性（Spatial Variations）生成共享密钥，因其克服了预分发密钥和密钥协商带来的不安全因素，得到了高度重视。具体来说，基于信道特征生成密钥利用了无线通信中衰落和噪声的客观存在，利用了信道特征参数的上下行信道的短时互易性、不同空间位置的空间差异性、不同时刻的时变性，将信道特征作为通信双方所共享的随机源，避免现有无线网络安全实现中的预密钥分发或密钥协商，具有轻量级、易操作等优点，特别是在缺乏密钥管理中心或可信第三方很难得到保证的自组织网、传感器网络中，具有广阔的应用市场。

2. 密钥构建典型框架

利用信道互易性建立密钥的方案大致都包括 3 个步骤：

① 信道特征参数评估阶段；

② 参数量化阶段，量化的基本思想是使用门限将采样值转化成二进制比特；

③ 信息调和及隐私加强阶段，如图 5-15 所示。

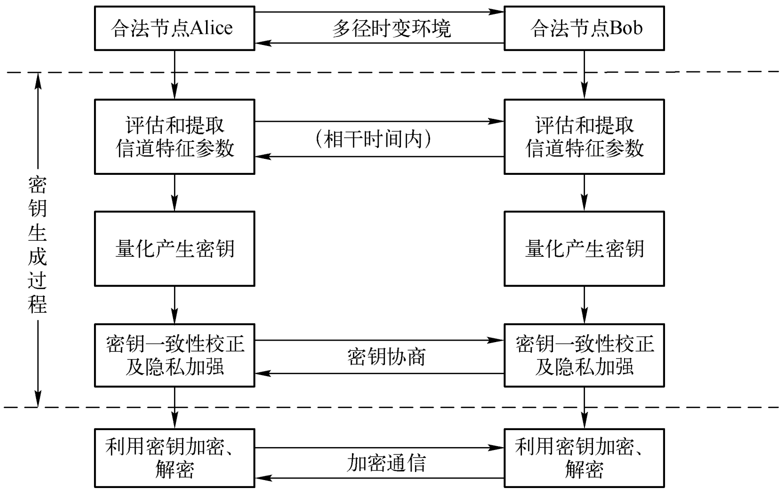


图 5-15　合法用户密钥建立流程图

密钥建立方案中，量化参数的提取和二进制比特的量化对于整个密钥建立方案的优劣发挥着决定性的作用；而对于信息调和及隐私加强则借鉴了其他领域的处理方法。

1）信道特征参数评估

信道特征参数评估技术的选择主要基于传输信号的带宽或信道的时域、频域特性。从传输信号带宽的角度来分，信道特征参数评估技术主要可以分为窄带和宽带信道探测技术两种，在特定时间间隔内两种技术中的发送者和接收者均在同一特定频率采样接收信号。

① 窄带信道探测技术：在信道衰落和延时作用下，到达接收机处的信号与原发射信号相比，在幅度和相位上都产生反映信道作用的变化。窄带信道探测技术正是基于这个现象，通常使用正弦信号作为探测信号，利用信号幅度和相位的变化来确定信道的特征。

② 宽带信道探测技术：宽带信道探测技术主要利用发送周期窄带脉冲信号或利用 PN 序列扩频的方式来获得较宽频带内的无线信道特性，常用的方法主要有滤波器技术、周期脉冲探测法、扩频滑动相关法及频域测量法等。

2）量化阶段

量化阶段是密钥建立过程中最重要的部分，通信双方依据特定的门限将信道状态信息量化成二进制比特。高性能的量化算法与信道的特征参数和量化门限的选择有关。

3）信息调和及隐私加强阶段

信息调和及隐私加强阶段是两个相对独立的研究领域，之所以把二者归在一个步骤，是因为在密钥构建中，二者联系紧密。由于不完美的互易性和随机噪声，通信双方初始独立量化的比特序列可能不完全一致，使用信息调和技术通信，交换反馈信息处理不一致的部分。信息调和的目的是使用最少量的信道信息更正或删除不一致的比特位。隐私放大技术是为了防止敌手利用信息调和阶段所泄露的信息破解密钥。

3. 密钥生成评价标准

物理层信道特征密钥构建性能评价指标主要有 3 个：密钥生成速率、密钥不一致率、随机性。

① 密钥生成速率：密钥生成速率 $R_K = N/T$（单位为 bit/s），其中 N 为时间 T 内生成比特流的数目，即被定义为单位时间内生成的密钥比特数目。R_K 反映着算法的效率，是密钥建立技术整体性能的关键指标。

② 密钥不一致率：密钥不一致率反映的是传输者和接收者得到初始的比特序列的不同性，反映着量化算法的性能。$P_K = 1 - (1 - P_e)^N$，其中 P_e 表示单个比特出错的概率，N 为比特流长度。P_K 反映着算法的强壮性，是评估量化算法性能的重要参数，较高的不一致率表明所采用的量化算法易受到随机信道噪声和不完美信道互易性的影响。

③ 随机性：随机性反映着密钥的安全性。在信息论中，使用熵的概念来表示随机变量的不确定性，用来评估共享密钥的安全强度。通常，高熵值表示随机变量的更大的随机性，敌手很难推导出高熵值的密钥。熵值的定义：

$$H_i = -p_0 \log p_0 - (1 - p_0) \log (1 - p_0)$$

$$H_{\text{total}} = \sum_{i=0}^{N} H_i$$

其中，N 代表密钥的总长度，p_0 代表变量的后验概率。实际操作中，随机性普遍使用 NIST（National Institute of Standards and Technology）测试。

5.4.3　一种无线物理层密钥生成机制

当前已经有很多利用无线信道特征生成密钥的方法。其中 Suhas Mathur 在“Radio - telepathy：extracting a secret key from an unauthenticated wireless channel”一文中提出一种基于接收信号强度（Received Signal Strength，RSS）和信道冲激响应（Channel Impulse Response，CIR）来对信道进行测量，从而提取密钥的方法。Mathur 提出的方法分为信息提取、量化、协商纠错与密钥生成几个步骤。

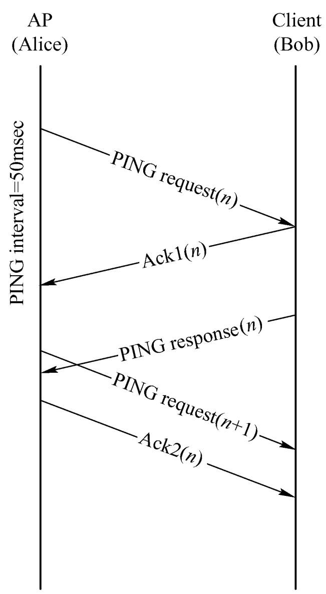


图 5-16　使用 PING 命令收集 RSS 值的时序图

为了保证 Alice 和 Bob 测量的信道特征有很强的相关性，Alice 和 Bob 测量信道的时间间隔要足够小。在该文中设计出带确认的 PING 协议，因为 PING 自带的 request - response 机制的时间间隔过长，Alice 和 Bob 收集的信道测量相关性低，不利于生成相同的密钥。Alice 发送 PING request，Bob 收到后，给 Alice 发送确认 Ack1；随后 Bob 对 Alice 的 PING request 回复 PING response，Alice 收到 response 消息后，给 Bob 发送确认 Ack2，如图 5-16 所示。其中 request 和 Ack1 及 response 和 Ack2 的序号是一致的，用来保证丢包重传及 Alice 和 Bob 收集的信道特征值之间的匹配。经过信息提取过程，Alice 和 Bob 之间已经共享了相关性很

强的信道特征值。量化阶段的主要工作是将这些信道特征值转换为 01 比特流。

量化阶段使用的量化器：

$$Q(x)=\begin{cases}0, x<q_- \\ 1, x>q_+\end{cases}$$

在平均值上下有两条量化线 q_- 和 q_+，将大于 q_+ 的值量化为 1，小于 q_- 的值量化为 0，落在 q_+ 和 q_- 之间的值舍弃，因为这一部分值很容易偏移到大于 q_+ 或小于 q_- 的区域中，造成误码率，如图 5-17 所示。经过量化阶段，无线信道特征值被转换成 01 比特流。但是这些比特流还不能作为密钥来使用。因为 Alice 和 Bob 生成的序列还不完全相同，需要通过密钥协商才能生成相同的比特流。这些比特流不相同的原因是 Alice 和 Bob 测量的信道特征值有较大的跳动，如 Alice 端某一值大于 q_+ 而在 Bob 端对应的值小于 q_-。密钥协商阶段的作用是将这些不同的位消除。密钥协商阶段使用连续 M 位相同比特输出一位来消除错误。具体就是连续 M 个 1 输出 1，连续 M 个 0 输出 0。

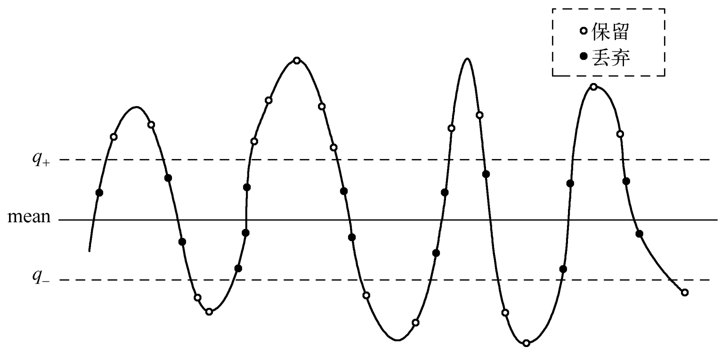


图 5-17　RSS 量化器

经过以上三步操作，Alice 和 Bob 生成相同的密钥。具体算法如下。

① Alice 测量自己的信道特征值，然后使用 $Q(x)$ 量化器对收集到的信道特征值进行量化，找出连续 M 个大于 q_+ 或小于 q_- 的特征值，记录下对应的中间位的索引，得到一个索引集。

② Alice 随机地选择索引集的子集，将该子集发送给 Bob。

③ 收到该索引集后，Bob 检测每个索引值的两边是否一共包含 $M-1$ 个大于 q_+ 或小于 q_- 的值。

④ 如果有索引值不满足上一步的条件，Bob 就将该索引值从索引集中剔除。Bob 将该索引集发送给 Alice。

⑤ Alice 和 Bob 根据这个索引集分别生成密钥。

其中步骤①属于信息提取阶段和量化阶段，步骤②、步骤④属于协商纠错和密钥生成阶段。

参考文献

[1] Shi E, Perrig A. Designing secure sensor networks. Wireless Communication Magazine, 2004, 11 (6): 38－43.

[2] Karlof C, Sastry N, Wagner D. TinySec: A link layer security architecture for wireless sensor networks. In: Proc. of the 2nd ACM Conf. on Embedded Networked Sensor Systems. New York: ACM Press, 2004: 162 – 175.
[3] Eschenauer L, Gligor V. A key management scheme for distributed sensor networks. In: Proc. of the 9th ACM Conf. on Computer and Communications Security. New York: ACM Press, 2002: 41 – 47.
[4] Chan H, Perrig A, Song D. Random key predistribution schemes for sensor networks. In: Proc. of the 2003 IEEE Symp. on Security and Privacy. Washington: IEEE Computer Society, 2003: 197 – 213.
[5] Du W, Deng J, Han YS, Varshney PK. A pairwise key pre – distribution scheme for wireless sensor networks. In: Proc. of the 10th ACM Conf. on Computer and Communications Security. New York: ACM Press, 2003: 42 – 51.
[6] Liu D, Ning P. Establishing pairwise keys in distributed sensor networks. In: Proc. of the 10th ACM Conf. on Computer and Communications Security. New York: ACM Press, 2003: 52 – 61.
[7] Liu D, Ning P. Location – Based pairwise key establishments for static sensor networks. In: Proc. of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks. New York: ACM Press, 2003: 72 – 82.
[8] Du W, Deng J, Han YS, Chen S, Varshney PK. A key management scheme for wireless sensor networks using deployment knowledge. In: Proc. of the IEEE INFOCOM. Piscataway: IEEE Press, 2004: 586–597.
[9] Huang D, Mehta M, Medhi D, Harn L. Location – Aware key management scheme for wireless sensor networks. In: Proc. of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks. New York: ACM Press, 2004: 29 – 42.
[10] Chan H, Perrig A. PIKE: Peer intermediaries for key establishment in sensor networks. In: Proc. of the IEEE INFOCOM 2005. Piscataway: IEEE Communication Society, 2005: 524 – 535.
[11] Camtepe SA, Yener B. Combinatorial design of key distribution mechanisms for wireless sensor networks. In: Proc. of theComputerSecurity—ESORICS. Berlin: Springer – Verlag, 2004: 293 – 308.
[12] Perrig A, Szewczyk R, Tygar J, Wen V, Culler D. SPINS: Security protocols for sensor networks. ACM Wireless Network, 2002, 8 (5): 521 – 534.
[13] Zhu S, Setia S, Jajodia S. LEAP: Efficient security mechanisms for large – scale distributed sensor networks. In: Proc. of the 10th ACM Conf. on Computer and Communications Security. New York: ACM Press, 2003: 62 – 72.
[14] Younis M, Ghumman K, Eltoweissy M. Location – Aware combinatorial key management scheme for clustered sensor networks. IEEE Trans. on Parallel and Distribution System, 2006, 17 (8): 865–882.
[15] Eltoweissy M, Moharrum M, Mukkamala R. Dynamic key management in sensor networks. IEEE Communications Magazine, 2006, 44 (4): 122 – 130.
[16] Berlekamp E R, McEliece R J, and Tiborg V. On the inherent intractability of certain coding problems. IEEE Trans. On Information Theory, 1978, 24 (3): 384 – 386.
[17] Blum A, Furst M, Kearns M, and Lipton R J. Cryptographic primitives based on hard learning problems. In Advances in Cryptology – CRYPTO'93, Vol. 773 of Lecture Notes in Computer Science, 1993: 278 – 291.
[18] Gilbert H, Robshaw M, Sibert H. An active attack against HB + ——Aprovably secure lightweight protocol. Electronics Letters, 2005, 41 (21): 1169.
[19] Bringer J, Chabanne H, Dottax E. HB + + : a lightweight authentication protocol secure against some attacks. Lyon: IEEE International Conference Oil Pervasive Services, Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing, 2006.
[20] Phan RC, Yen S. Amplifying Side – Channel Attacks with Techniques from Block Cipher Cryptanalysis. Tarragona: Proceedings of the 7th Smart Card Research and Advanced Application IFIP Conference, 2006.
[21] Eschenauer L, Gligor V D. A key – management scheme for distributed sensor networks: proceedings of the 9th

ACM Conference on Computer and Communications Security. 2002：41 －47.

[22] Zhu S，Setia S，Jajodia S. LEAP：efficient security mechanisms for large －scale distributed sensor networks：proceedings of the 10th ACM Conference on Computer and Communication Security . Washington D C，2003：62 －72.

[23] Chan H，Perrig A，Song D. Random key predistribution schemes for sensor networks. In：Proc. of the 2003 IEEE Symp. on Security and Privacy. Washington：IEEE Computer Society，2003：197 －213.

[24] Younis M，Ghumman K，Eltoweissy . ocation －Aware ombinatorial key management scheme for clustered sensor networks. IEEE Trans. on Parallel and Distribution System，2006，17（8）：865－882.

[25] Eltoweissy M，Moharrum M，Mukkamala R. Dynamic key management in sensor networks. IEEE Communications Magazine，2006，44（4）：122 －130.

[26] 李琳，王汝传，姜波，黄海平 . 无线传感器网络层簇式密钥管理方案的研究 . 电子与信息学报，2006，28（12）：2394 －2397.

[27] 黄鑫阳 . 杨明无线传感器网络密钥管理研究综述 . 计算机应用研究，2007，24（3）：10 －15.

[28] S. Mathur，W. Trappe，N. B. Mandayam. Radio －telepathy：extracting a secret key from an unauthenticated wireless channel. In ACM MOBICOM Conference，Sept. 2008.

第 6 章　感知层数据安全传输技术

数据传输是网络通信的基础，是任何网络构建之前首先要解决的问题。物联网感知层数据传输面临窃听、碰撞、干扰、路由攻击等多种安全威胁，本章仍以典型的 RFID 系统和无线传感器网络为例，给出物联网安全数据传输的一些研究成果。同时在本章最后还会向读者介绍近年来比较热门的网络编码技术，以及网络编码在数据安全传输中的具体应用。

6.1　RFID 系统安全通信技术

RFID 系统中，阅读器与电子标签通信过程中会存在很多干扰，其中最主要的干扰就是信道噪声和多卡操作，这些干扰会使传输的信号发生畸变，从而导致传输的错误。要提高数字传输系统的可靠性，就要采用差错控制技术，对可能或已经出现的差错进行控制。当网络中存在多个阅读器和标签时，采用适当的数据传输防碰撞技术可减少通信冲突，提高效率。此外，RFID 作为物联网重要的信息获取手段，在向物联网提供相关信息时，必须向公众确保个人隐私保护。

6.1.1　RFID 差错控制技术

通信信道上总有噪声存在，由于噪声会对有用信息进行干扰，因此数据传输会出现差错。在数字通信中，常常采用差错控制来确保接收数据的完整性和准确性。

1. 差错的产生

通信过程中的差错大致可以分为两类：一类是由热噪声引起的随机错误，另一类是由冲突噪声引起的突发错误。突发错误影响局部，而随机错误影响全局。

1）随机错误

热噪声引起的差错是一种随机差错，亦即某个码元的出错具有独立性，与前后码元无关。传输随机错误的信道称为无记忆信道或随机信道。

2）突发错误

突发错误是由冲突噪声引起的。冲突噪声是由短暂原因造成的，如电机的启动或停止，电器设备的放弧等。冲突噪声引起的差错是成群的，它们之间有相关性，其差错维持时间称为突发错误的长度。传输突发错误的信道称为有记忆信道或突发信道。

2. 差错控制的基本方法

一般利用编码方法对传输中产生的差错进行控制。数据信息在向信道发送之前，先按照某种关系增加监督码元，即附加上一定的冗余位，并利用监督码元去发现传输中发生的错误，这个过程称为差错控制编码过程。接收端收到该码字后，检查信息和附加冗余位之间的

关系，以检查传输过程中是否有差错发生，这个过程称为检验过程。

1）差错检验编码

差错检验编码可以分为检验码和纠错码。

① 检错码：能自动发现差错的编码。

② 纠错码：不仅能发现错误，而且能自动纠正差错的编码。

2）差错控制方法

差错控制方法主要分为两类，一类是反馈纠错 ARQ，另一类是前向纠错 FEC。在这两类的基础上，又派生出一种混合纠错。对于不同类型的信道应该采用不同的差错控制技术。

（1）反馈纠错

反馈纠错（ARQ）方式是能发现传输差错的编码方法。这种方法在发送端加入少量监督码元，在接收端根据编码规则对收到的信号进行检查，当发现有错码时，即向发送端发送询问信号，要求重发，直到信息被完全接受为止。

ARQ 协议的优点是它非常简单，因而被广泛地应用在分组交换网络中。ARQ 协议的缺点是需要接收方发送应答信号 ACK，这样增加了网络的负担也影响了传输速率。重复发送数据包来纠正错误的方法也严重地影响了它的传输速率。

（2）前向纠错

前向纠错（FEC）是一种在单向通信系统中控制传输错误的技术，通过连同数据发送额外的信息进行错误恢复，以降低误码率（Bit Error Rate，BER）。采用前项纠错方式时，不需要反馈信道，也无须反复重发而延误传输时间，这对实时传输有利。但是，FEC 方式的纠错设备比较复杂。

（3）混合纠错

混合纠错的方式是综合使用反馈纠错和前向纠错的方法。当少量纠错时，采用前向纠错的方法，在接收端自动纠正。当差错较严重，超出自行纠正能力时，采用反馈纠错的方法，向发送端发出询问的信号，要求重发。

3）纠错原理及常见校验码

为了使信源编码具有检错和纠错的能力，应按一定的规则在信源编码的基础上增加一些冗余码元（又称监督码元），使这些冗余码元与被传送信息码元之间建立一定的关系。在收信端根据信息码元与监督码元的特定关系，可以实现检错或纠错。

（1）奇偶校验码

奇偶校验码是一种最简单的线性分组检错编码方式，通过增加冗余位使得码字中 1 的个数恒为奇数或偶数。在实际使用时它又可分为垂直奇偶校验、水平奇偶校验和水平垂直奇偶校验等几种。

① 垂直奇偶校验。

垂直奇偶校验又称纵向奇偶校验，它是将要发送的整个信息块分为定长 p 位的若干段（如 q 段），每段后面按“1”的个数为奇数或偶数的规律加上一位奇偶位 r，如图 6-1 所示。

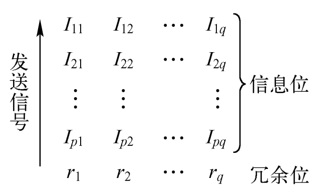


图 6-1　垂直奇偶校验

垂直奇偶校验方法能检测出每列中的所有奇数位的错，但检测不出偶数位的错。对于突发错误，奇数位错与偶数位

错的发生概率接近于相等，因而对差错的漏检率接近于 1/2。

② 水平奇偶校验。

为了降低对突发错误的漏检率，可以采用水平奇偶校验方法。水平奇偶校验又称横向奇偶校验，它是对各个信息段的相应位横向进行编码，产生一个奇偶校验冗余位。水平奇偶校验不但可以检测出各段同一位上的奇数位错，而且还能检测出突发长度小于等于 p 的所有突发错误，因为按发送顺序从图 6-2 中可见，突发长度小于等于 p 的突发错误必然分布在不同的行中，且每行一位，所以可以检查出差错，它的漏检率比垂直奇偶校验方法低。但是实现水平奇偶校验码时，不论是采用硬件还是软件方法，都不能在发送过程中边产生奇偶校验冗余位边插入发送，而必须等待要发送的全部信息块到齐后，才能计算冗余位，也就是一定要使用数据缓冲器，因此它的编码和检测实现起来都要复杂一些。

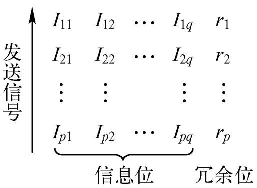


图 6-2　水平奇偶校验

③ 水平垂直奇偶校验。

同时进行水平奇偶校验和垂直奇偶校验就构成了水平垂直奇偶校验，也称纵横奇偶校验。

（2）CRC 校验

CRC 即循环冗余校验码，是数据通信领域中最常用的一种差错校验码，其特征是信息字段和校验字段的长度可以任意选定。

① CRC 的基本原理。

CRC 校验是在 K 位信息码后再拼接 R 位校验码，整个编码长度为 N 位，因此，这种编码又叫 (N,K) 码。对于一个给定的 (N,K) 码，可以证明存在一个最高次幂为 $N-K=R$ 的多项式 $G(x)$。根据 $G(x)$ 可以生成 K 位信息的校验码，而 $G(x)$ 叫作这个 CRC 码的生成多项式。

② CRC 码的生成步骤。

- 将 x 的最高幂次为 R 的生成多项式 $G(x)$ 转换成对应的 $R+1$ 位二进制数。
- 将信息码左移 R 位，相当于对应的信息多项式 $G(x)\times 2^R$。
- 用生成多项式（二进制数）对信息码做除，得到 R 位的余数。
- 将余数拼到信息码左移后空出的位置，得到完整的 CRC 码。

③ CRC 检验过程。

设编码前的原始信息多项式为 $P(x)$，$P(x)$ 的最高幂次加 1 等于 K；生成多项式为 $G(x)$，$G(x)$ 的最高幂次等于 r；CRC 多项式为 $R(x)$。

- 发送方编码方法：将 $P(x)$ 乘以 x^r（即对应的二进制码序列左移 r 位），再除以 $G(x)$，所得余式即为 $R(x)$，用公式表示为 $T(x)=x^rP(x)+R(x)$。
- 接收方解码方法：将 $T(x)$ 除以 $G(x)$，如果余数为 0，则说明传输中无错误发生，否则说明传输有误。

标准的 CRC 生成多项式见表 6-1。

表 6-1　标准的 CRC 生成多项式

名　　称	生成多项式	简　　记	应 用 举 例
CRC－4	x^4+x+1	3	ITU G. 704
CRC－12	$x^{12}+x^{11}+x^3+x+1$	80F	
CRC－16	$x^{15}+x^{15}+x^2+1$	07	IBM SDLC

续表

名　称	生成多项式	简　记	应用举例
CRC－ITU	$x^{16}+x^{12}+x^5+1$	8005	ISO HDLC，ITU X. 25，V. 34/V. 41/V. 42
CRC－32	$x^{32}+x^{26}+x^{23}+\cdots+x^2+x+1$	04C11DB7	ZIP，RARIEEE 1394，PPP－FCS
CRC－32c	$x^{32}+x^{28}+x^{27}+\cdots+x^8+x^6+1$	1EDC6F41	SCTP

6.1.2　RFID 数据传输防碰撞技术

随着 RFID 技术的发展，阅读器通信距离不断增加，其识别区域的面积也越来越大，进而导致可能出现多个标签同时出现在阅读器的识别范围之内。在多个阅读器和多个标签的 RFID 系统中，主要存在两种形式的冲突：一种是同一标签同时收到不同阅读器发出的信号，这称为阅读器碰撞；另一种是一个阅读器同时收到多个不同标签返回的信号，信号将产生叠加而导致阅读器不能正常解析标签发送的信号，这称为标签碰撞。第一种情况在实际应用中是可以避免的，而第二种情况在实际应用中却很容易出现。这样当有两个或两个以上的标签同时给阅读器发送数据时，就会出现标签之间的冲突，因此必须采取适当的通信方式来解决这个问题，物联网标签防冲突技术可以通过硬件、软件两种途径来实现。解决该通信冲突问题的方法被称为防碰撞技术。

1. 无线网络中的防碰撞技术

一般来说，无线网络中有四种解决碰撞的方法，即时空分多址（Space Division Multiple Access，SDMA）、码分多址（Code Division Multiple Access，CDMA）、频分多址（Frequency Division Multiple Access，FDMA）和时分多址（Time Division Multiple Access，TDMA）。下面简要介绍四种方法的工作原理。

1）SDMA

这种技术是利用空间分割构成不同的信道，可以实现频率的重复使用，充分利用频率资源。举例来说，在一颗卫星上使用多个天线，各个天线的波束射向地球表面的不同区域。地面上不同地区的地球站，它们在同一时间即使使用相同的频率进行工作，也不会形成干扰。

SDMA 实现的核心技术是智能天线的应用，理想情况下它要求天线给每个用户分配一个点波束；这样根据用户的空间位置就可以区分每个用户的无线信号，换句话说，处于不同位置的用户可以在同一时间使用同一频率和同一码型而不会相互干扰。实际上，SDMA 通常都不是独立使用的，而是与其他多址方式如 FDMA、TDMA 和 CDMA 等结合使用；也就是说，对于处于同一波束内的不同用户再用这些多址方式加以区分，增加了容量。

2）CDMA

在扩频通信技术上发展起来的 CDMA 技术，是为现代移动通信网所要求的大容量、高质量、综合业务、软切换、国际漫游等要求而设计的一种移动通信技术。CDMA 技术的原理是基于扩频技术，即将需要传送的具有一定信号带宽的信息数据，用一个带宽远大于信号带宽的高速伪随机码进行调制，使原数据信号的带宽被扩展，再经载波调制并发送出去。接收端使用完全相同的伪随机码，与接收的带宽信号作相关处理，把宽带信号换成原信息数据的

窄带信号即解扩，以实现信息通信。

CDMA 移动通信网由扩频、多址接入、蜂窝组网和频率复用等几种技术结合而成，含有频域、时域和码域三维信号处理的一种协作，因此它具有抗干扰性好、抗多径衰落、保密安全性高、同频率可在多个小区内重复使用、容量和质量之间可做权衡取舍等属性。同时，它也存在以下缺点：来自非同步 CDMA 网中不同的用户的扩频序列不完全正交，从而引起多址干扰；由于使用相同的载频，许多用户共用一个信道，强信号对弱信号有着明显的抑制作用，从而产生“远——近”效应，影响用户通信质量。

3）FDMA

把信道频带分割为若干更窄的互不相交的频带（称为子频带），把每个子频带分给一个用户专用（称为地址），这种技术被称为“频分多址”技术。所有信道都可以作为单信号被扩大、控制，并转换为频带传送至目的地，该技术的主要优点在于经济实用。

FDMA 模拟传输是效率最低的网络，这主要体现在模拟信道每次只能供一个用户使用，使得带宽得不到充分利用。此外，FDMA 信道大于通常需要的特定数字压缩信道，且沉默状态下 FDMA 信道也是浪费的。模拟信号对噪声较为敏感，并且额外噪声不能被过滤出去。

4）TDMA

TDMA 是通信技术中的基本多址技术之一，在 2G（GSM）移动通信系统中多被采用，用于卫星通信和光纤通信的多址技术中。

时分多址是把时间分割成周期性的帧（Frame），每一个帧再分割成若干个时隙向基站发送信号，在满足定时和同步的条件下，基站可以分别在各时隙中接收到各移动终端的信号而不混扰。同时，基站发向多个移动终端的信号都按顺序安排在预定的时隙中传输，各移动终端只要在指定的时隙内接收，就能在合路的信号中把发给它的信号区分并接收下来。

2. RFID 防碰撞技术

RFID 防碰撞问题与计算机网络媒介访问层中的网络碰撞实质上是一样的，但由于 RFID 系统尤其是标签硬件的能力有限，如标签没有冲突检测功能、标签之间不能互相通信、所有的冲突仲裁（或冲突判断）都需要由阅读器完成、标签存储容量和计算能力有限等，传统网络中的很多计算量过于繁重、硬件能力要求高的算法将不再适用。基于 SDMA、FDMA、CDMA 的标签防碰撞算法理论上可以解决部分有源的标签防碰撞问题，但是目前大多数使用的是无源标签，受其功率的限制，采用上述方案会使得标签和阅读器的设计复杂、系统成本高。从 RFID 系统的复杂性及成本考虑，TDMA 是最有实际用价值也最常见的一类防碰撞方法。TDMA 防碰撞协议分为 Tree 协议和 Aloha 协议，Tree 协议是一种确定性机制，它能识别阅读器感知区域内的所有标签，但时延较长；而 Aloha 协议采用随机机制，得益于其简单且易于实现而被广泛应用于 RFID 标准中。

1）纯 Aloha 防碰撞算法

纯 Aloha 防碰撞算法主要采用标签先发言（tag－talk－first）的方式，即标签一旦进入阅读器的工作范围获得能量后，便向阅读器主动发送自身的序列号。在标签向阅读器发送数据的过程中，如果有其他电子标签也同时向阅读器发送数据，此时阅读器接收到的信号就会

产生重叠，导致阅读器无法正确识别和读取数据。阅读器通过检测并判断接收到的信号是否发送碰撞，一旦碰撞产生，阅读器就向标签发送指令使得标签停止数据传输；标签接收到指令后，随机延迟一段时间重新发送数据。

由于各个标签等待的时间是随机的，因此一定程度上避开了标签数据的碰撞，但显然碰撞的次数与通信业务量的大小有关。纯 Aloha 防碰撞算法的主要特点就是各个标签发射时间不需要同步，是完全随机的，实现简单，但只能在标签比较少的情况下效果才好。纯 Aloha 防碰撞算法简单易实现，但是信道利用率并不高。分析表明，纯 Aloha 防碰撞算法的信道吞吐率 S 与帧产生率 G（单位时间内发送的信息帧数量）之间的关系为

$$S = G\mathrm{e}^{-2G}$$

例如，当 $G = 0.5$ 时，信道吞吐率 $S = 18.4\%$。

2）时隙 Aloha 防碰撞算法

时隙 Aloha 防碰撞算法是基于纯 Aloha 协议提出的。阅读器将一帧划分为多个时隙，每个时隙大于或等于标签标识符发送时间长度。在所有标签和阅读器取得同步基础上，规定标签仅能在每个时隙开始时才能发送数据。时隙 Aloha 防碰撞算法流程如下。

① 当阅读器向标签发送请求命令时会同步向所有标签广播时隙个数 L（即帧长）。

② 标签会在帧长范围内随机地选择一个时隙响应阅读器的命令并发送自身的信息，若一个时隙中只有一个标签返回信息则称为成功时隙，没有标签返回信息则称为空时隙，有两个或更多的标签返回信息则称为碰撞时隙，记一帧中空时隙数为 C_0，成功时隙数为 C_1，碰撞时隙数为 C_k。发生碰撞的相关标签会在下一帧继续向阅读器发送数据。

③ 算法根据前一帧的反馈值 C_0、C_1、C_k 采用标签估算方法来估算阅读器范围内未读标签数量，并根据此调整下一帧时隙数，得到一个使系统识别效率最高的时隙数 L'。

④ 读写器向标签广播新的时隙数，直至所有标签被识别完。

在时隙 Aloha 防碰撞算法中，信道的利用率有所提高。分析表明，时隙 Aloha 防碰撞算法的信道吞吐量 S 与帧产生率 G 之间的关系为

$$S = G\mathrm{e}^{-G}$$

当 $G = 0.5$ 时，信道吞吐率 $S = 30.3\%$。

该算法的优点在于逻辑简单，电路设计简单，所需内存少，而且在帧内只随机发送一次，这样能够进一步降低冲突的概念。时隙 Aloha 防碰撞算法是 RFID 系统中最常见的一种基于 Aloha 的防碰撞算法。

3）基于二进制树的防碰撞算法

二进制树算法目前应用非常广泛，之所以称为“二进制树”，是因为在算法执行过程中阅读器要多次发送命令给电子标签，每次命令都把标签分为两组，多次分组后最终得到唯一的一个标签。其基本思想是按照递归的工作方式将阅读器工作区域内的标签不断地划分为 p 个子集（$p > 1$），再对某个子集进行同样划分，这样所划分子集内的标签数量越来越少，直至某子集内的标签数目为 1，实现阅读器成功识别标签。当某个子集内的标签读取完毕，阅读器采用回溯的方式处理其他等待读取的标签。这样就可以将标签的分组过程看成全部标签根据分组方案从根节点向叶节点逐层分流的过程，只有叶节点的标签能被成功读取。

划分子集的算法有两种，一种是让标签随机选择所属集合，这种算法称为随机二进制树

算法；另一种是按照标签的标志符号进行划分，这种算法称为查询二进制树算法。

（1）随机二进制树算法（见表 6-2）

表 6-2　二进制树算法过程

阅读器发送序号	请求 1111 1111	第一次操作	请求 1011 1111	第二次操作	请求 1010 1111	第三次操作	选择 1010 0011	读写
标签应答		1×1×001×		101×001×		不响应		
标签 A		10110010		10110010		不响应		
标签 B		10100011		10100011		10100011		10100011
标签 C		10110011		10110011		不响应		
标签 D		11100011		不响应		不响应		

算法规定每个标签都要有唯一的序列号用以区别不同的标签，且各标签必须同步发送序列号，即在同一时刻开始传送它们的序列号。只有这样阅读器才能检测出碰撞位。算法基本步骤如下。

- 首先阅读器向电子标签发送一个最大序列号，所有标签序列号小于或等于该值的标签向阅读器回送其序列号。
- 由于标签序列号的唯一性，当小于最大序列号的标签数量不小于两个时，必然发生碰撞，阅读器检测到碰撞后，将最大序列号中对应的碰撞起始位置为 0，低于该位者不变，高于该位者置 1。
- 阅读器将处理后的序列号发送给标签，标签序列号与该值比较，小于或等于该值的标签，将自身的序列号返回给阅读器。
- 循环这个过程，就可以选出一个序列号最小的标签，然后阅读器与该标签进行正常通信，通信结束以后发出命令使该标签进入休眠状态，即除非重新上电，否则不再响应阅读器的命令请求。也就是说，下一次阅读器再发最大序列号时，该标签不再响应。
- 重复上述 4 步，即可按照序列号从小到大依次识别出各个标签，算法的过程示意见表 6-2。

（2）查询二叉树算法。

查询二叉树算法是一个无状态协议，不需要标签内部维持任何状态，标签只需要根据阅读器广播的标志符前缀做比较即可。其算法原理：阅读器发送长度为 k 的 prefix；标签 ID 中前 $k$bit 与 prefix 匹配的标签反馈第 $(k+1)$bit 至最后 1bit。如果阅读器收到标签 ID 发生碰撞，再分别将 prefix 加“1”和“0”，作为新的 prefix 发送出去。如果没有碰撞，就表明一个标签被识别了。

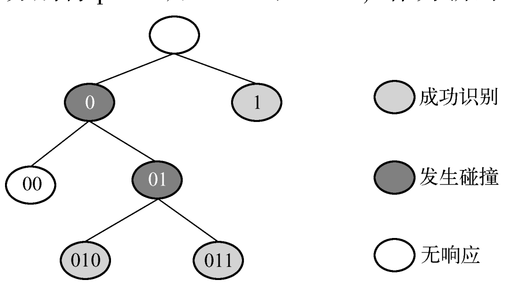


图 6-3　查询二叉树算法

以图 6-3 为例，设标签的三个 ID 分别为 010、011、100，阅读器的查询序列首先置为 0、1，阅读器先发送序列 0 进行查询，发生碰撞，此时将序列置为 00、01，再次分别发送，序列 00 没有响应，序列 01 发生碰撞，将序列置为 010、011，成功识别。回溯到序列 1，只有标签 100 响应，成功识别。

6.2　传感器网络安全路由技术

6.2.1　无线传感器网络路由协议概述

1. 无线传感器网络路由与传统路由的典型差别

在无线传感器网络研究初期，人们一度认为成熟的 Internet 路由技术加上 Ad – hoc 网络路由机制对无线传感器网络路由设计是足够充分的，但深入的研究表明，无线传感器网络有着与传统网络明显不同的技术要求，主要包括以下几方面。

- 能量有限。传统的路由协议在选择最优路径时，很少考虑节点的能量消耗问题，而无线传感器网络中节点的能量有限，延长整个网络的生存期成为传感器网络路由协议设计的重要目标，因此需要考虑节点的能量消耗及网络能量均衡使用的问题。
- 基于局部拓扑信息。无线传感器网络为了节省通信能量，通常需要建立多条通信路径，而节点有限的存储资源和计算资源，使得节点不能存储大量路由信息，不能进行太复杂的路由计算。在节点只能获取局部拓扑信息和资源有限的情况下，如何实现简单高效的路由机制是无线传感器网络的一个基本问题。
- 以数据为中心。传统的路由协议通常以地址作为节点的标识和路由的依据，无线传感器网络中的大量节点随机部署，所关注的是检测区域的感知数据，而不是具体哪个节点获取的信息，不依赖于全网唯一的标识。传感器网络通常包含多个传感器节点到少数汇聚节点的数据流，按照对传感数据的需求、数据通信模式和流向等，以数据为中心形成信息的转发路径。
- 应用相关。传感器网络具有很强的应用相关性，不同应用的路由协议差别很大，没有一个通用的路由协议适用于所有网络。

因此，包括 Ad – hoc 网络在内的传统网络路由技术都无法直接应用于无线传感器网络，针对传感器网络的应用需求及自身特点，需要设计专门的路由协议。目前无线传感器网络研究领域已经提出了很多路由协议，如洪泛路由、数据为中心的路由、层次路由、基于地理位置的路由等，但是很多路由机制都没有考虑路由的安全问题。

2. 无线传感器网络路由安全威胁

无线传感器网络节点数量大、部署环境恶劣、维护困难，相比 Ad – hoc 网络更容易受到攻击，比较常见的攻击主要有虚假路由信息攻击、选择性转发攻击、污水池攻击、女巫攻击、虫洞攻击、Hello 泛洪攻击、确认攻击等。需要注意的是，传感器网络路由的攻击并不局限于这些，很多攻击是以上两个或多个相互结合的新型攻击，如女巫攻击可以和虫洞攻击相结合对地理位置路由机制进行攻击。路由协议的安全威胁见表 6–3。

表 6–3　路由协议的安全威胁

路由协议	安全威胁
TinyOS 信标	虚假路由、选择性转发、污水池、女巫、虫洞、Hello 泛洪
定向扩散	虚假路由、选择性转发、污水池、女巫、虫洞、Hello 泛洪
地理位置路由	虚假路由、选择性转发、女巫

续表

路由协议	安全威胁
最低成本转发	虚假路由、选择性转发、污水池、女巫、虫洞、Hello 泛洪
谣传路由	虚假路由、选择性转发、污水池、女巫、虫洞
能量节约的拓扑维护（SPAN、GAF、CEC）	虚假路由、女巫、Hello 泛洪
聚簇路由协议（LEACH、TEEN）	选择性转发、Hello 泛洪

1）虚假路由信息攻击

最直接的攻击方法是针对节点间路由交换信息进行攻击。通过伪造、篡改或重放路由信息，攻击者可以形成路由回路，恶意吸引或拒绝网络流量，产生错误的路由信息，将网络分离，增加端到端延迟。

2）选择性转发攻击

多跳网络能够正常运行的一个基本前提就是所有参与节点都能诚实地转发所有收到的信息。但在选择转发攻击中，恶意节点并不诚实，它可能拒绝转发特定的消息，或者直接丢掉一些数据包，使得这些数据包无法继续被传播。其中一种简单的攻击方式就是黑洞攻击，一个节点将它能见到的所有数据包全部都丢弃，形成路由黑洞。但在黑洞攻击中，攻击者可能很容易被周围其他邻居发现，并认为该节点已经无法正常工作，而选择其他节点作为转发节点。

3）污水池攻击

污水池攻击有时也称黑洞攻击，攻击者的目标是吸引从一个区域来的几乎所有的数据流通过一个已经受到入侵的节点，使它看起来对周围基于一定路由算法的节点更具有吸引力。通过正式存在的或虚假的通往已被入侵节点的高性能路由，攻击者周围的每个节点很可能会把转发目的地的数据包交给攻击者传输，并且向各自相邻的节点传播这个颇具吸引力的路由消息。

4）女巫攻击

在女巫攻击中，一个节点在网络中的其他节点面前具有多个不同的身份。女巫攻击能够明显地降低路由方案对于诸如分布式存储、分散和多径路由、拓扑结构保持的容错能力。女巫攻击还给基于位置信息的路由协议造成了很大的威胁。对于位置敏感的路由，为了高效地为基于地理位置标识的数据包选路，通常要求节点与它们的邻居交换坐标信息。一个节点对于它的邻居们来说只有唯一的一组坐标才是合理的，但是通过使用女巫攻击，攻击者可以同时处在不同的坐标上。

5）虫洞攻击

攻击者通过一个很小延迟链路形成隧道，利用隧道在网络不同部分间传输收到的消息，并且在网络的不同部分重放这些消息。最简单的例子就是，两个节点串通合谋进行攻击。一个恶意节点在基站附近，另一个相距较远，这个节点声称自己和基站附近节点可以建立低时延高带宽的链路，以及吸引其他节点把其数据包发往这里。虫洞攻击很可能与选择性地转发或女巫攻击相结合。当它与女巫攻击相结合时，通常很难探测出这些攻击。

6）Hello 泛洪攻击

一些路由协议需要节点不断发送 Hello 包，以声称自己为邻居，但当一个较强的恶意节

点以大功率广播 Hello 包时，收到此包的节点会认为这个恶意节点是它们的邻居。在以后的路由中，这些节点可能会使用恶意节点的路径，使得网络不能正常运行。使用 Hello 泛洪攻击的攻击者不必具有构造合法数据流的能力。它只以足够让网络中的所有节点都能收到消息的功率，重新广播系统开销数据包就可以了。因此，Hello 泛洪攻击可以想象成一个单向广播虫洞攻击。

3. 无线传感器网络路由协议设计要求

1）能量高效性

无线传感器网络路由协议不仅要选择能量消耗小的信息传输路径，而且要从整个网络的能量均衡消耗的角度进行路由选择。传感器节点的资源有限，传感器网络的路由机制要能够简单而高效地实现信息传输。

2）安全性

无线传感器路由安全旨在保障网络路由信息的可获得性、路由信息的完整性和报文的可靠路由。由于节点的移动性，使其自身的资源和能力受限，并且网络缺乏有效的物理保护，这些都使得无线自组织网络路由机制面临多种安全威胁。

3）可扩展性

在无线传感器中，因为节点失败、新节点加入及节点移动等，都会使网络拓扑结构动态发生变化，这就要求路由机制具有可扩展性，能够适应网络结构的变化。

4）鲁棒性

能量用尽或环境因素造成传感器节点的失败，周围环境影响无线链路的通信质量及无线链路本身的缺点等，这些无线传感器网络的不可靠特性要求路由机制具有一定的容错能力。

5）快速收敛性

传感器网络的拓扑结构动态变化，节点能量和通信带宽等资源有限，因此要求路由机制能够快速收敛，以适应网络拓扑的动态变化，减少通信协议开销，提高信息传输的效率。

4. 无线传感器网络路由协议分类

传统的路由协议无法适应无线传感器网络的需要，因此必须选择或设计适用于无线传感器网络环境特点的路由协议。经过多年的研究，许多协议方案相继被提出来，这些路由协议根据不同的分类标准有不同的分类方式，如图 6-4 所示。

① 根据路由协议所依据的基本路由算法不同，无线传感器网络路由协议可以分为：基于链路状态的路由协议（如 OLSR、STAR、TBRPF 等）、基于距离矢量的路由协议（如 DS-DV）、源路由协议（如 DSR）和反向链路协议（如 TORA）。

② 根据路由建立的方式不同，可分为：先应式路由协议（如 DSDV、OLSR、TBRPF 等）、按需路由协议（如 DSR、AODV、ABR 等）和混合式路由协议（如 ZRP、CEDAR 等）。

③ 根据路由协议所依据的网络逻辑结构的不同，可分为：平面结构的路由协议（如 AODV、DSR、DSDV 等）和分层结构的路由协议（如 CGSR、HSR、VBS 等）。

④ 根据路由协议所适用的网络规模不同，可分为：中、小规模的路由协议（如 AODV、

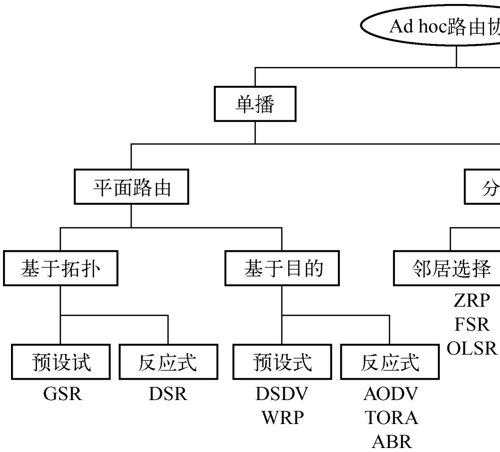


图 6-4　无线传感器网络路由协议分类

DSR 等）和大规模的路由协议（如 FSR、CGSR、HSR、LANMAR 等）。

⑤ 根据接收业务数据的目的节点个数的不同，可分为：单播路由协议（如 DSDV、LAR、AODV 等）和多播路由协议（如 MAODV、ODMRP、AMROUTE 等）。

另外，还有一些路由协议使用特殊的机制或完成特殊的功能，如通过地理位置信息来改善路由协议的性能或根据地理位置信息进行数据转发，称为基于地理位置信息的路由协议（如 LAR、GPSR 等）。某些路由协议具备提供 QoS 保证的能力，称为 QoS 路由协议（如 CE-DAR、SRL、QoS－OLSR 等）。还有一些路由协议专门设计用来节省网络功率消耗，延长网络寿命，称为功率感知的路由协议（如 COMPOW、PARO、GAF 等）。

6.2.2　传感器网络信息协商路由协议（SPINS）

鉴于传感器网络面临的诸多威胁，必须要为传感器网络设计合适的安全防护机制，保护整个网络安全通信。SPINS 安全协议框架是最早的无线传感网络安全框架之一，包括 SNEP（Secure Network Encryption Protocol）和 uTESLA（micro Timed Efficient Streaming Loss－tolerant Authentication Protocol）两个部分，在 4.3 节已经部分介绍过。这里重点介绍 SPINS 的数据协商过程。

SPINS 中提出了元数据（Meta－data，即描述传感器节点采集的数据属性的数据）的概念，它是对具体数据的抽象描述，小于节点实际采集的数据，因此有相应的数据请求时才发送数据信息。在传输或接收数据之前，通过采用资源自适应机制，每个节点都必须检查各自可用的能量状况，如果处于低能量水平，必须中断一些操作，如充当数据中转的角色、停止数据转发操作等。

SPINS 采用三次握手协议来实现数据的交互，协议运行过程中使用三种报文数据：ADV、REQ 和 DATA。ADV 用于数据的广播，当某一个节点有数据可以共享时，可以用 ADV 数据包通知其邻节点，该数据包保护元数据；REQ 用于请求发送数据，当某一个收到 ADV 的节点希望接收 DATA 数据包时，发送 REQ 数据包；DATA 为原始数据包，里面装载原始感知数据，同时包含元数据的头部。

如图 6-5 所示，SPINS 的工作过程大致分为三个步骤。

步骤 1：当一个传感器节点（源节点）有新数据需要传输时，使用 ADV 数据包（包含元数据）对其所有的邻居进行广播，等待接收 REQ 响应消息。

步骤 2：当邻居节点收到这个广播消息时，首先检查其自身的能量值是否低于设定阈值，若能量充足则检查是否已经接收过或请求过该广播的数据；若请求接收该数据，且其拥有足够的能量，就发送 REQ 消息请求接收新的数据。若已经接收过或请求过该数据或能量低于设定阈值，则不响应。

步骤 3：若源节点没有收到 REQ 消息，则进入结束状态，等待下一次数据传输；若源节点收到 REQ 消息，则发送数据 DATA 给请求节点。接收节点若不是汇聚节点，在接收数据后就变为源节点，返回到步骤 1，继续执行。这个过程重复下去，直到数据被传输到汇聚节点。

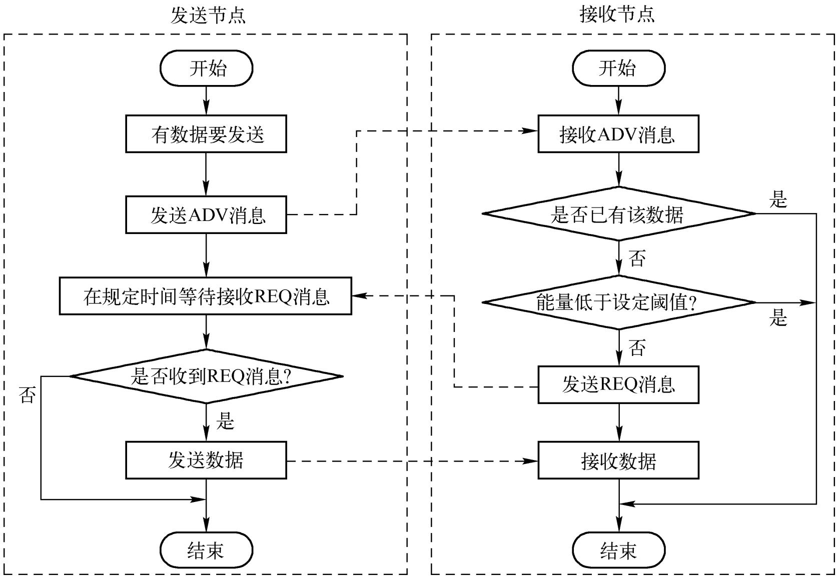


图 6-5　SPINS 发送节点和接收节点工作示意图

SPINS 定义的是一个协议框架，在使用时还要考虑很多具体的实现问题。例如，使用什么样的加密、鉴别、认证、单向密钥生成算法和随机数发生器，如何在有限资源内融合各种算法以达到最高效率等。

1. 加密算法的选择

传感器网络的计算能力和存储能力要求只能用对称密钥的块加密算法。AES 标准的 Rijindael 算法需要 800 字节的查找表；改进的 Rijindael 虽然加密速度快了 100 倍，但是查找表增加到了 10K 字节。这对于只有几千字节的传感器节点来说是不合适的。DES 算法需要 512 项的 S 盒和 256 项的变换盒，对于存储有限的无线传感器节点而言直接使用是十分困难的。RC5 算法比较合适。RC5 算法简单高效，不需要很大的表支持。最重要的是该算法是可定制的加密算法。对于要求不同，节点能力不同的应用可以选择不同的定值参数，非常灵活和方便。

2. 消息认证算法的选择

信息完整性和新鲜性保证都需要消息认证算法。消息认证算法通常使用单向散列函数。

目前最常用的单项散列函数有 MD5、SHA、CBC – MAC 等。

3. 密钥生成算法

基站和节点之间共享的主密钥对是在网络部署之前就确定了的，在通信过程中要谨慎选择由主密钥对生成的通信密钥和认证广播密钥对的生成算法。通信密钥是通过单向散列函数作用于主密钥上产生的。对于密钥更新问题 SPINS 中没有过多考虑，为了保证通信密钥的前向保密，一种方法是通过以后安全通道协商一个完全不同的密钥，这种方式的通信开销大；另一种方法就是通过一个更为巧妙的方法利用单向散列函数和通信密钥生成不相关、免协商的密钥。

4. 随机数发生器

真正的随机数计算机很难生成，实际中通常利用伪随机数发生器来实现。伪随机数发生器往往通过一个函数连续产生一串看起来无序的数据串。伪随机数产生的随机数因为由一个固定的函数产生，一旦输入相同，将产生一样的随机序列，所以在利用伪随机数产生随机数列时，一定要定期修改种子或使用生成较慢的真随机数作为伪随机数的种子。种子改变，随机数序列的顺序就会发生改变，从而有更好的随机效果。

6.2.3 INSENSE 入侵容忍路由协议

INSENSE 协议（Intrusion – tolerant routing protocol for wireless sensor networks）是一种可以在一定程度上容忍攻击的路由协议。该协议能够建立有效的树形路由，并适合于无线传感器网络中非对称结构和资源受限的特点。INSENSE 的主要思想是在路由协议中加入容侵机制，该方法主要考虑如何承受攻击而非对抗攻击。

1. INSENSE 的工作原理

采用 uTESLA 中的单项序列和加密的消息鉴别码算法，及时发现入侵者，用基站寻找建立路由和用基站向每个节点发送请求信息来增强网络的健壮性，使网络具有一定的对抗攻击能力和自我修复能力。另外，多径选择策略也可以在一定程度上增强网络的容侵性。

2. INSENSE 的三个原则

① 该协议用来及时发现入侵者，由于入侵行为的检测非常耗时，所以该协议采用了容侵的路由机制。在路由初始化阶段，每个节点保留了多条到基站的路由。

② 针对节点能源受限的问题，用基站来寻找和建立路由，基站的责任就是计算路由。

- 基站向每个节点发送请求信息，收到请求信息的节点再向它的邻居节点转发请求信息，并把向它发送请求信息的节点标记为邻居节点，每个节点只转发一次信息。
- 节点向基站发送应答信息。
- 基站计算出每个节点到基站的路由，当基站计算出所有的路由信息后，它会向节点发送一个前向转发表，在这个转发表中包含了节点到基站的几条不同路由。

③ 该协议用于入侵者未检测出的情况下，如何减少入侵者所造成的破坏。在本协议中，基站对节点采用单向认证，减少入侵者提供的错误路由信息、洪泛反馈信息或不断向邻居节点转发请求信息等所造成的破坏。而且，只允许节点与基站间进行通信，因此入侵者不能与

邻居节点直接通信。

3. INSENSE 的缺点

INSENSE 存在一定的缺点，如没有把节点的能耗问题作为设计的目标，而且入侵行为的检测非常耗时，过多地依赖于基站等。另外，INSENSE 还不能很好地承受无线传感器中所面临的各种安全攻击。

6.2.4　协作式安全路由协议

协作式安全路由协议的主要思路是：利用安全协议来抵抗常见的安全攻击，并将声誉机制引入安全协议，通过声誉机制来检测节点，依据一定的原则对节点的行为予以评价，进而检测出不良节点，并给予相应的处置，激励节点合作，保证路由的性能和网络的安全。

1. 基于 BETA 分布的声誉机制

1）直接声誉的计算

节点 i 用 BETA 分布作为其保存的关于节点 j 的直接声誉的先验分布，即 $F_{i,j}$，其形式为 (α_n,β_n)，初始化为（1,1）。节点 i 一旦观察到 j 进行了一项新的协作活动，它就更新 $F_{i,j}$，并用改进的贝叶斯估算法来更新 α_n、β_n 这两个参数。

2）综合声誉的计算

综合声誉 $R_{i,j}$表示的是节点 i 对节点 j 的网络行为的综合看法，包括 j 的协作性好坏，是否正确转发包，是否正确执行路由协议等。用改进的贝叶斯估算法来处理 $R_{i,j}$的更新，$R_{i,j}$更新是基于以下两类事件的：

- 节点 i 更新其关于节点 j 的直接声誉 $F_{i,j}$；
- 网络中其他节点 k 发布了关于 j 的声誉报告 $F_{k,j}$，i 通过一定的判定规则判定 $F_{k,j}$是可信的，并用于更新 $R_{i,j}$。

3）诚实值的计算与更新

节点的诚实分布函数 $T_{i,k}$和诚实值 $T^*_{i,j}$的计算与更新也使用改进的贝叶斯估算法。$T_{i,k}=(\lambda_n,\upsilon_n)$，其初始值为(1,1)。设节点 i 当前保存的关于节点 k 的诚实分布为$(\lambda_{n-1},\upsilon_{n-1})$。当节点 i 每次收到来自 k 的间接声誉 $F_{k,j}(\alpha_k,\beta_k)$时，它首先会对 $F_{k,j}$进行一致性测试，测试 $F_{k,j}$是否与 $R_{i,j}(\chi_{n-1},\eta_{n-1})$一致，测试的方法就是计算$|E(R_{i,j})-E(F_{k,j})|$，用 ψ_n 表示测试结果。若$|E(R_{i,j})-E(F_{k,j})|<d$，则认定节点 k 发布的信息可靠，$F_{k,j}$可信，$\psi_n=1$；否则令 $\psi_n=0$。

4）节点分类与行为决策

节点 i 要进行与节点 j 有关的网络行为决策之前，它先计算 $R^*_{i,j}=E(R_{i,j})$与 $T^*_{i,j}=E(T_{i,j})$。节点 i 可以根据 $R^*_{i,j}$，$T^*_{i,j}$将节点分类：

$$\begin{cases}\text{不良节点}: R^*_{i,j}<t_\theta\\ \text{良性节点}: R^*_{i,j}\geqslant t_\theta\end{cases}$$

$$\begin{cases}\text{不诚实节点}:\text{若 } T^*_{i,j}<t_t\\ \text{诚实节点}:\text{若 } T^*_{i,j}\geqslant t_t\end{cases}$$

对于良性节点，网络中的其他节点才考虑接受其通信请求，这样抵御了恶意节点的非法通信请求，拒绝接受恶意节点的异常数据，从而在一定程度上保障了节点数据的可能性。

2. 协作式安全路由设计

在协议中每个节点保存着其他节点的声誉信息。这些声誉信息包括 $F_{i,j}$，$R_{i,j}$，$T_{i,j}$，$F^*_{i,j}$，$R^*_{i,j}$，$T^*_{i,j}$。初始化时 $F_{i,j}=R_{i,j}=T_{i,j}=(1,1)$，故 $F^*_{i,j}=R^*_{i,j}=T^*_{i,j}=0.5$。

1）路由发现过程

源节点 A 需要传送数据到目的节点 X，A 首先发起一轮路由查找过程，向其邻居节点广播一个 RDP 包。

$$A \to broadcast: [RDP, IP_X, Cert_A, N_A, t_A]K_{A-},\ Cert_A$$

如果中间节点 B 认为源节点 A 是一个良性节点且有合作传送的意向，在验证了 A 的身份与 RDP 的完整性后，将目的节点地址 IP_X、前一跳节点地址 IP_A 信息存储下来，然后用自己的私钥 K_{B-} 对收到的包进行签名并附上自己的证书继续广播转发。

$$B \to broadcast: [[RDP, IP_X, Cert_A, N_A, t_A]K_{A-}, Cert_{A-}]K_{B-},\ Cert_B$$

其邻居节点 C 收到 RCP 后，首先根据节点 B 的证书来验证它的签名及 RDP 包的来源。若正确，节点 C 记录目的节点地址 IP_X、前跳节点地址 IP_B，然后去掉 RDP 中 B 的签名再用自己 C 的私钥对内容进行签名，加上自己的证书继续转发包，直到发现目的节点 X。

$$C \to broadcast: [[RDP, IP_X, Cert_A, N_A, t_A]K_{A-}, Cert_{A-}]K_{C-}, Cert_C$$

2）数据发送

节点 A 需要发送数据包到目的节点 X。首先节点 A 根据路由表中的记录，选择声誉值最高的节点 B，然后单播发送数据包至节点 B。节点 B 根据节点 A 的声誉值，将其插入发送队列相应位置中，若节点 A 的声誉值较低，则将节点 A 的数据包插到发送队列的最后，若节点 A 的声誉值最高，就将节点 A 的数据包插到发送队列的前面。节点 B 在接收到节点 A 传送过来的数据包后启动一个定时器并保存一个数据包的副本，然后把数据包发送给声誉值最高的下一条节点（假设为节点 C）。

目的节点 X 接收到正确的数据后会向源节点发送 DACK 包以确定数据传送正确。中间节点 B 在定时器过期之前从其下游节点收到一个 DACK 报告，则从缓存中删除副本，并按照声誉机制提升下游节点的声誉。

3）声誉评价阶段

在数据传输成功并收到 DACK 确认包的情况下，路由上的节点认为其下一跳节点进行了一次成功的协作活动，则其按照声誉机制更新路由下一条节点的直接声誉分布与综合声誉分布及其他数据。如果在某个设定的时间内未收到 DACK 包，就认为路由上存在一个不良节点，这时路由上的节点将通过采用声誉机制综合自身及邻居节点对下一跳节点的声誉进行评价。

6.3 网络编码技术在数据传输中的应用

网络编码是由 Ahlswede 等人于 2000 年针对有线网络中组播问题的研究提出的。它通过允许中间节点不仅对接收到的包进行存储转发，而且对接收到的多个包进行编码融合，使网

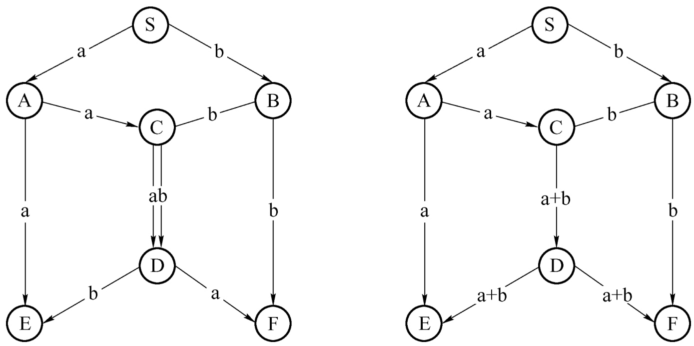
络多播容量达到最大流理论的极限值。

6.3.1　网络编码的基本原理及分类

在现有的通信网络中，网络的中继节点只是将接收到的消息数据存储并转发，而不对数据作其他处理。但是，网络编码理论的出现打破了这种常规。网络编码的核心思想是允许网络的中继节点对接收到的消息数据进行编码，如模 2 加、有限域上的运算等，而不是仅仅复制转发，进而达到最大流传输。

1. 网络编码概述

图 6-6 所示的蝴蝶网络是网络编码实现多播最大容量的例子。源节点 A 和 B 分别组播 1 比特到目的节点 E 和 F，假设各链路的容量为 1。图 6-6（a）中采用的是传统路由方法，即节点 C 一次只能传送 1 比特到节点 D，节点 D 也只能传送 1 比特到节点 E 和 F，节点 C 和 D 之间的链路不得不使用了两次，节点 E 和 F 总共收到 3 比特，平均速率是 1.5 比特/单位时间。图 6-6（b）采用网络编码方法，节点 C 对输入信息流进行编码，将编码的结果 a 和 b 的异或值（模 2 和）传送到节点 D，再传送给节点 E 和 F。节点 E 根据自身已收到的信息 a 和 a+b，可以解码出 b。同样，节点 F 也能解码出完整的信息，这样所能达到的平均速率是 2 比特/单位时间。在本例中，采用网络编码使得每条链路只使用了一次，这样既使得网络负载比较均衡，又节省了传输次数，同时减小了网络时延，增大了网络吞吐量。网络编码主要应用于组播传输，类比于路由方式，其数据分发也包括以下两个基本步骤。



（a）传统路由　　（b）网络编码

图 6-6　蝴蝶网络

1）选择最佳编码子图

类似于选择最佳路由，目的是选择数据从源到目的节点的最佳分发路径。网络编码网络模型图通常用一有向无环图 $G=(V,E)$ 来表示，其中 V 是节点集，E 是边集。$V_s=\{s_1,s_2,\cdots,s_\sigma\}$ 是信源节点集合，$V_D=\{t_1,t_2,\cdots,t_\rho\}$ 表示信宿集。对任意的节点 $u,v\in V$，若存在从 v 到 u 的有向信道，则用有向边 $e=(v,u)$ 表示此信道，且 u 称为 e 的头节点，v 称为 e 的尾节点，分别记为 $v=T(e)$，$u=H(e)$，边 e 传输的消息数据记为 $Y(e)$。对任意的节点 v，$\Gamma_I=\{e\in E:H(e)=v\}$ 和 $\Gamma_o=\{e\in E:T(e)=v\}$ 分别为节点 v 的入边集和出边集。

2）确定编码方案

在编码子图中确定每个节点对收到的数据包的编码和转发方式。

2. 网络编码分类

1）线性网络编码和非线性网络编码

全体边对应的编码函数的集合称为网络的编码，它描述了该网络的传输方案。若所有编码函数都是线性函数，则网络编码称为线性的，否则称为非线性的。由于线性编码函数较容易实现，所以大部分研究工作都集中在线性网络编码，对非线性编码函数的研究较少。

2）确定型线性网络编码和随机线性网络编码

若知道通信网络的部分或全部信息（如网络的拓扑结构、相邻节点的状态、网络开销等），则在消息传输前就可以确定消息的传输方案，计算出通信网络中各信道的全局网络编码。这就是确定型线性网络编码方法。确定型线性网络编码需要事先知道网络的拓扑结构信息，根据所掌握的网络的相关信息来确定编码方式，以实现最优的信息传输。为了克服确定型线性网络编码在实际应用中的局限性，研究者提出了随机线性网络编码方法，该方法不需要事先已知网络拓扑，中间节点编码时只需要对来自不同链路的包进行随机线性组合，就可以使接收节点以一定概率成功译码。译码的概率与编码所使用的有限域 F_q 的大小有关。F_q 越大，译码概率越高，但是编译码的计算复杂度也越大。随机网络编码将在下一节重点介绍。

3. 网络编码优点

- 网络编码在无线网络中的应用可以提高网络的吞吐量，尤其是组播吞吐量；
- 可以减少数据包的传播次数，降低无线发送能耗；
- 采用随机网络编码，即使网络部分节点或链路失效，最终在目的节点仍然能恢复原始数据，增强网络的容错性和鲁棒性；
- 无须复杂的加密算法，采用网络编码就可以提高网络的安全性，减少了计算复杂度。

6.3.2 随机网络编码技术

确定性网络编码 DNC（Determine Network Coding）的编码构造方法都是基于已知网络拓扑信息的集中式方法，在应用中具有一定的局限性。针对这个问题，Ho、Medard 等人在 2003 年提出了不用网络拓扑信息的随机网络编码方法。在该方法中节点只对来自不同链路的数据包进行随机的线性组合处理（节点输出是输入的随机线性组合，组合的系数从一个有限域内随机选取），就可以使得接收节点以一定的概率成功解码。研究结果表明随着接收节点数量与编码运算基于的有限域大小的比值趋近于 0，该算法中接收节点能够成功解码的概率将趋近于 1。

总的来看，集中式的网络编码方法需要根据全局拓扑状况来给每个节点分配编码系数，虽然这些方法使得编码运算所需的有限域不会太大，但是网络拓扑的变化将导致全部编码系数重新分配，因此其可扩展性较差。随机网络编码的提出拓宽了网络编码的适用场景，使得网络编码不再局限于确定的网络拓扑和集中式的算法，具有更大的实用性。

1. 随机网络编码的编码原理

在随机编码策略下，对于所有除了接收节点外的节点，我们在一个足够大的有限域上随机选择它们输入链路到输出链路的映射。每一个节点的映射选取是相互独立的。对于随机网

络编码，每一个接收节点只需要知道其每一个输入符号对应的信源随机过程的总的线性组合，即每一个输入的边的全局编码向量。

在分组网络中，一条链路上传输的符号被分组传输，则流过一条边 e 的符号$y(e)$ 可以划分为向量：

$$y(e)=[y_1(e),y_2(e),\cdots,y_N(e)]$$

同时，流过信源 s 的第 i 条输出边的信源符号 x_i 可以分组为向量：

$$x_i=[x_{i,1},x_{i,2},\cdots,x_{i,N}]$$

对于任何接收节点有：

$$\begin{bmatrix} y(e_1) \\ \vdots \\ y(e_h) \end{bmatrix}=\begin{bmatrix} y_1(e_1) & \cdots & y_N(e_1) \\ \vdots & \ddots & \vdots \\ y_1(e_h) & \cdots & y_N(e_h) \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_h \end{bmatrix}=G_t\begin{bmatrix} x_{1,1} & \cdots & x_{1,N} \\ \vdots & \ddots & \vdots \\ x_{h,1} & \cdots & x_{h,N} \end{bmatrix}$$

为了使接收节点可以获得每一条输入边对应的全局编码向量信息，给第 i 个向量 x_i 的头部添加一个单位向量，该单位向量的第 i 个元素为 1。这样信息通过网络的传输，对于每一个接收节点，其每条输入边上传输的信息向量的头部就是改变对应的全局编码向量。

$$\begin{bmatrix} g_1(e_1) & \cdots & g_N(e_1) & y_1(e_1) & \cdots & y_N(e_1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_1(e_h) & \cdots & g_N(e_h) & y_1(e_h) & \cdots & y_N(e_h) \end{bmatrix}=G_t\begin{bmatrix} 1 & \cdots & 0 & x_{1,1} & \cdots & x_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & x_{h,1} & \cdots & x_{h,N} \end{bmatrix}$$

这样接收节点只要通过高斯消元法就可以恢复出信源发送的信息向量 $x_1,x_2,\cdots,x_h$。每一分组附加发送一个编码向量的开销对系统而言是很小的。通过该方式，接收节点在译码时可以不知道多播网络的拓扑结构或编码函数，网络中的节点及边的添加或删除对接收节点的译码不会产生影响。

2. 随机网络编码译码原理

设 $g(e)$为边 e 对应的全局编码向量，接收节点 t 的 h 条输入边 $e_1,e_2,\cdots,e_h$ 对应的输入符号为：

$$\begin{bmatrix} y(e_1) \\ y(e_2) \\ \cdots \\ y(e_h) \end{bmatrix}=\begin{bmatrix} g_1(e_1) & \cdots & g_h(e_1) \\ \vdots & \ddots & \vdots \\ g_1(e_h) & \cdots & g_h(e_h) \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_h \end{bmatrix}=G_t\begin{bmatrix} x_1 \\ \vdots \\ x_h \end{bmatrix}$$

要使接收节点恢复出信源符号 $x_1,x_2,\cdots,x_h$，只要全局编码向量 $g(e_1),g(e_2),\cdots,g(e_h)$ 组成的矩阵 G_t 满秩。当编码的符号域足够大时，即便随机选择局部编码向量也可保证 G_t 满秩。

对于一个信源相互独立或线性相关的可行的多播链接问题，如果一个网络码字的一些或所有编码系数都独立并且在一个有限域上均匀选取（一些节点的编码系数可以是固定的，只要这些取值是可行的），则接收者就可以最小概率$(1-d/q)^v$ 恢复出信息序列，这里的 q 为编码符号域的大小，d 是接收节点的个数，v 是随机选取编码系数的链路的个数。

因为编码是在整个码长 u 上进行的，所以随机网络编码的复杂度随着编码符号域的增大而增大。但是译码失败的概率却与域的大小成反比。如果随机编码的符号域的大小至少为

$|E|/\delta$，则相应的编码对于任何给定的接收者至少以概率 $1-\delta$ 可逆。当符号域的大小至少为 $|E||T|/\delta$ 时，编码对于所有接收者至少以概率 $1-\delta$ 同时可逆。

3. 随机网络编码优势

① 随机网络编码是一种分布式的网络编码方式。网络编码时，它不用知道整个网络的拓扑结构，就可以进行编码。

② 随机网络编码在每条链路上随机选取编码向量进行数据包编码，攻击节点很难通过监听链路传输的有限数据包进行原始数据包恢复，因此对于能力受限的攻击节点的攻击有着较好的防御效果。

③ 编码符号域足够大时，可保证接收节点以很大的概率恢复出信源的信息。当符号域 $|F|=2^{16}$ 时，理论上任意一个接收节点可以至少以概率 0.996 成功获取信源发送的信息。在实际环境中，域的大小为 2^8 就足够了。

④ 随机网络编码可适用于随机的网络结构。对于网络节点和链路时变的网络，随机网络编码可以利用整个网络的剩余容量来获得网络的最佳容量，提高网络多播的鲁棒性。

6.3.3　COPE：一种实际的编码路由协议

早期的网络编码大多关注于编码方式的研究，多进行理论推导以提高多播流量，而在具体的应用研究方面乏善可陈。COPE 是第一个实用的网络编码机制，它起初在无线网状网（Wireless Mesh Network，WMN）提出了一个有效的单播传输架构。

1. COPE 原理

COPE 是由 Katti 等人提出的一个实际的无线网络编码系统。COPE 利用了无线链路的广播特性和协议栈设计的灵活性，提出了机会侦听（opportunistic listening）和机会编码（opportunistic coding）的概念：每个无线节点向周围的邻居汇报自己已得到的数据，而每个节点在转发数据时，通过周围邻居已获得数据的情况选择网络编码的方案。

1）COPE 关键技术

（1）机会侦听（opportunistic listening）

无线网络具有广播的特性。当无线节点配置有全向天线时，无线节点将会侦听到很多数据包，在某种情况下可以用来增加吞吐量。COPE 协议正是利用了无线的广播特性，将所有节点的网卡设为混杂模式，存储在 T 时间（默认 0.5s）内侦听到的所有包。此外，COPE 利用无线网的广播特性，采用了伪广播方式，其实质是在 802.11 单播包上捎带一个 XOR 头，里面指明了该编码包的所有下一跳节点，而该单播包的 MAC 目的地设为下一跳集中的一个即可。由于 COPE 网络中的所有节点都是混杂模式，所以它们可以听到 MAC 地址不属于自己的包。当收到包时，该节点检查 XOR 头，如果自己在下一跳集中，则处理该包；否则将该包放入存储池中，当作机遇侦听得到的包。由于所有的包都是单播，所以可以利用 802.11 的冲突检测和回退机制。

此外，所有节点会广播自己的接收报告（reception reports），告诉邻居自己所存的包。为了减少不必要的能量消耗，接收报告一般是附属在数据包上的，如果没有数据发送，也会

周期性地广播控制包。

（2）机会编码（opportunistic coding）

机会编码的关键问题是，把哪些包 XOR 到一起可以达到最大的吞吐量。原则上认为，为了最大化吞吐量，一个节点在发送时，需要最大化一次传输中本地包的数量，并且要保证下一跳集中的每个节点都能够解码。

从不同的数据流发送到同一个中间节点的数据包有多种编码方式。但是有的编码方式下，不需要的编码数据包将被发送到根本不能解码的领域，造成容量的损失。一般认为，最为简单可达的高效编码方案应满足如下规则：一个节点要将 $p_1, p_2, \cdots, p_n$ 这 n 个数据包混在一起，传给 $r_1, r_2, \cdots, r_n n$ 个下一跳，当且仅当 r_i 有 $n-1$ 个 $p_j(j! = i)$。

这个法则保证了每个下一跳集中的节点都能解码。当一个节点有机会发送时，它选择满足上述法则的最大的 n，进而最大化编码的收益。

（3）获得邻居信息

如前所述，通过接收报告，可以通知邻居自己拥有的本地包信息。当严重拥塞时，接收报告也可能丢失；当流量很小时，接收报告的到来可能很迟，以至于附近的节点已经进行了次优的编码选择。因此，一个节点不能仅依赖于接收报告，而要猜测邻居所拥有的包。当 COPE 缺乏绝对信息时（指收到邻居的接收报告），它使用上一跳和下一跳之间的投递概率来作为该邻居拥有这个包的概率。当 COPE 猜测失败时，会导致一些下一跳无法解码。相应的包会被重传，并且可能和另外一组包编码到一起。

2）COPE 增益

吞吐量的增益来自编码机会，而编码机会又由流量模式决定。COPE 的编码增益主要来自以下两个方面。

（1）编码增益

定义：无编码时的传输次数与有编码时最少的传输次数的比值，主要用来衡量编码的效果。

在没有机会侦听的情况下，最大的编码增益是 2，并且是可达的。实际中观测到的收益比理论的要小，原因有编码机会的缺失、包头开销（COPE 要在 IP 前面加上许多自己的信息）、介质中的损失率。尽管如此，COPE 可以将传输的信息率增加到大于介质的比特率，这使得即使当介质被充分使用时，COPE 依然仍带来收益。

（2）编码 + MAC 增益

实际实验中，会发现增益超过了理论上的编码增益。这是因为 COPE 编码和 MAC 之间的交互产生了有益的副作用，所以叫作编码 + MAC 增益。

以图 6-7 为例，在没有编码的情况下，要使网络达到最大的吞吐量，Relay 的发送速率（发送机会）需要是 Alice 和 Bob 的两倍。但 802.11 的 MAC 协议为了保证公平性，三个节点的发送机会是相等的。这就使得在 Relay 中，有一半的包无法发出去。在使用 COPE 的情况下，Relay 可以同时发送来自 Alice 和 Bob 的包，相比没有 COPE 时，吞吐量变成了原来的两倍。

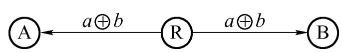


图 6-7　编码中继转发

编码 + MAC 增益假设所有节点都持续有数据发送，但受 MAC 分配的带宽所限制。当存在单点瓶颈时，编码 + MAC 增益可以由该节点带 COPE 的排出率（draining rate，指一个中继节点把输出队列中的包发送出去的速率）除以不带 COPE 的排出率来计算。

2. COPE 实现

1）包编码原则

为了建立编码方案，需要明确如下原则。

① 不延迟本应该发送的包。当无线信道空闲时，取出输出队列中的第一个包，检查是否有其他可以 XOR 在一起的包。如果没有编码机会，COPE 不会等待可能的匹配包，而是直接发送。

② 尽量将长度相近的包 XOR 在一起。因为把短包 XOR 到长包中时，会额外补零，导致带宽利用率不高。

③ COPE 不会把同样下一跳的包 XOR 在一起。否则，下一跳无法解码。

2）包格式

如图 6-8 所示，数据包包括以下三个部分。

① 该编码包中各本地包的信息，包括编码包中原始包的数目 ENCODED_NUM、包编号 PKT_ID、下一跳 NEXT_HOP。

② 接收报告，包括报告的总数目 REPORT_NUM、源 IPSRC_IP、上一个 IP 序列最后听到的数据包的 IP 序号 LAST_PKT、位图 Bit Map。

③ Ack 捎带，包含一个局部的 SEQ，用于局部 Ack 和重传。每个 Ack 条目包含（邻居标识 NEIGHBOR，最后一个局部 Ack 编号 LAST_ACK，位图 Bit Map）。

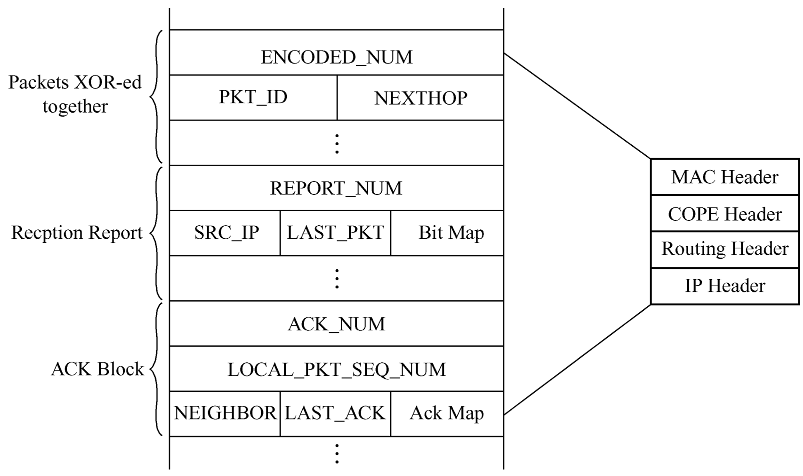


图 6-8　COPE 包头

为了保证 ACK 传输时消耗最少的费用，COPE 使用积累的 ACK。积累的 ACK 具有更好的鲁棒性以对抗数据包丢失。每一个节点都包含 16 位的邻居计数器 Neighbour_Seqno_Counter。一个 COPE 可以在一个条目中实现多个序列号（IP 序列号或 Ack 号）。COPE 使用（上一编号，位图）的表达方式。由于存储池中包的序列号和 Ack 的序列号都有数据局部性，所以使用上一编号 + 位图可以压缩信息。以 Ack 的一个条目为例：{A,50,01111111}，这个 Ack 表示该节点从 A 收到了 50 号及 43 ~ 49 号，但 42 号还是缺失的。

虽然 COPE 有累积 Ack 的机制，但并不保证链路级的可靠性。每个节点都只重传少数几次（默认为 2），然后放弃。

3）包编码算法

先从输出队列中取出队头的包，判断是大包还是小包。在相同尺寸的虚队列中去搜索可以 XOR 在一起的包。为防止乱序，COPE 只考察每个虚队列的队头。当一种尺寸的队列都搜索完后，才会搜索另一种尺寸的队列。在这些限制条件下，COPE 只用考察 $2M$ 个包头就行了（M 是邻居的数量）。

另一个主要问题是乱序。TCP 会把乱序当作网络拥塞的信号。COPE 按照输出队列的顺序发包，并且搜索虚队列时只考察队首，这极大程度地规避了乱序的问题。尽管如此，由于 COPE 分了大小包队列，所以乱序仍然可能发生。例如，当需要重传时，可能是由于组装 XOR 包时进行了错误的猜测，导致某节点无法解码。COPE 有一个重排模块可以把 TCP 分组排序后再递交上层。

最后，要使得编码包所到达的节点有最大的解码可能性。COPE 必须决定某个本地包是否可以被邻居解码。当该邻居是这个本地包的上一跳，或者收到了从该邻居发来的接收报告时，答案是肯定的。当上述条件不成立时，COPE 利用路由协议计算出的投递概率来估计某下一跳含有一个本地包的概率。COPE 确保下一跳集中每个节点都能解码的概率足够大。

某下一跳能解码的概率是它含有除它所需本地包外其他所有包的概率乘积：$PD = P1 * P2 * \cdots * P(n-1)$。假设 COPE 已经决定了编码 $n-1$ 个包，当决定第 n 个包时，它考察 n 个下一跳的 PD，是否都大于一个阈值 G（默认为 0.8）。这样保证了所有下一跳节点都有至少 G 的概率解码。最后，为了公平性，在虚队列间循环时使用随机顺序。

4）包解码算法

解码算法比较简单。每个节点都维护有一个存储池，拥有所侦听到的包的拷贝，并且以包编号为 key 组织成一个 hash 表。收到编码包时，只需要在存储池中搜索 $n-1$ 个已经侦听到的包，即可完成解码。存储池会周期地（几秒）做垃圾回收。

6.3.4　一种基于网络编码的延迟容忍移动传感器网络广播数据传输机制

广播作为一种重要的网络数据传输方式，主要用于网络全局信息的发布，如通过广播可以实现网络功能、任务的更新，通过广播可以进行全局消息的传输、网络共享密钥的分发等。相对于单播模式，广播能够有效降低网络通信开销、数据传输延迟，因此备受关注。DTMSN 网络中，基站不但需要对节点采集的信息进行收集，也需要根据实际情况向网络中移动的节点进行数据广播，因此针对广播数据传输机制的研究也成为 DTMSN 数据传输的重要内容。

DTMSN 中由于节点间的机会连接使得传统的基于簇、协商、多点中继的广播机制无法适用，因此 DTMSN 广播机制必须重新设计。当前，DTMSN 网络广播机制主要有直接传输（DT）、泛洪（Flood）、k-邻居（k-neighbor）等。

DT 是 DTMSN 中最基本的数据传输策略，适用于单播、多播及广播等多种数据传输需求。其基本思想是节点在网络中只与基站进行通信，而节点之间没有信息交互。当基站利用 DT 机制实现广播数据的传输时，实际上是利用多次单播来实现广播的目的。由于 DT 机制中节点只接收基站的广播数据，因此节点通信开销非常小，但由于节点并不是随时都可以和基站发生连接的，因此要想使所有节点都收到广播数据，时延一般较大，具体时延与连接发生的频率和数据传输的数量有关。

泛洪也是一种基本的数据传输策略，与 DT 相反，泛洪机制中基站会把需要广播的数据转发给所有能够与其通信的其他节点，而接收到广播数据的节点也会将数据泛洪给所有能够与其发生连接的其他传感器节点。泛洪机制通过大量的数据转发达到了有效降低广播延迟的目的，但由此也导致广播通信开销较大。为了降低通信开销，可控的泛洪机制、Spray - Wait 等机制相继被提出，这些机制主要是通过控制消息副本数量来降低网络的通信开销的，虽然这些机制能够在一定程度上提高节点能量的有效性，但带来的缺点是增加了广播传输延迟。

为了降低广播过程中的通信开销，有人提出了 k - 邻居（k - neighbors）机制，该机制的基本思想是充分利用无线信道的广播特性，只有在节点最少存在 k 个邻居的情况下才进行广播数据传输，从而避免了泛洪机制中向每个节点都进行数据传输所带来的通信开销。可见，k 越大，节点广播通信开销就越小。但由于节点需要有 k 个以上连接才能进行数据的广播传输，因此广播延迟较高。

近几年，网络编码在 DTMSN 网络中的应用研究也相继出现，这里给出一种基于网络编码的数据广播传输机制（Netcoding - based Broadcasting Transmission scheme，NBT）。该机制中传感器网络的基站将原始数据分批进行传输，且传输给不同传感器节点的数据采用的编码向量互不相关。由于基站对原始数据包进行了编码，因此传感器节点间具有相同编码数据包的可能性大大降低，从而减小了节点间的数据相关度，有效降低了广播时延。

1. 基础知识

1）节点数据相关度

定义 6.1：任意两节点 i、j 的数据相关度是指 i、j 具有的相同数据包的数量与它们全部数据包（不包括重复）数量的比，用 ρ_{ij} 表示。

例如，节点 i 有 8 个数据包，j 有 10 个数据包，其中相同的数据包数量为 6，则节点 i、j 的数据相关度为 $\rho_{ij}=6/(8+10-6)=0.5$。ρ_{ij} 越大，表示节点具有相同数据包的比例越高。

节点数据相关度能够有效衡量节点存储数据相同的概率，在 DTMSN 网络广播数据传输中，如果节点数据相关度高，则说明节点之间能够相互交换的数据越少，彼此从对方处能够获取的新的信息也就越少，要想取得全部广播信息的时间也就越长；相反，如果节点数据相关度低，则说明节点间能够相互交换的数据越多，彼此获取的信息量也就越大，获得全部广播消息的时间也就越短。

2）节点广播收益

为了衡量源节点在对多个目标节点进行广播数据传输过程中能够节省的通信开销，此处提出节点广播收益的概念。

定义 6.2：假设在没有通信差错的情况下，节点广播收益是指节点采用广播方式向其通信范围内其他节点传输广播数据时比其采用单播方式能够节省的通信开销的量。这里用节省消息的数量来衡量广播收益的大小，用 G 表示。

假设网络中有 A,B,C,D,E,F 6 个节点，基站需要广播的数据为 $a\sim h$ 共 8 个。例如，节点 B 能够提供给 C、D 的报文分别为 c,d,h 和 g,h，如果采用单播方式则需要传输 5 个报文，而采用广播方式则只传输 c,d,g,h 4 个报文，因此广播收益为 1。从表 6-4 中可以看出，节

点 B、C 如果利用自身与多个其他节点存在连接这一优势，采用广播的方式进行数据传输则能够有效降低数据传输的通信开销。

表 6-4　节点广播收益

节　点	单　　播	广　　播	收　　益
A	A→C：{*b*,*c*,*d*}	A→C：{*b*,*c*,*d*}	0
B	B→C：{*c*,*d*,*h*}，B→D：{*g*,*h*}	B→C,D：{*c*,*d*,*g*,*h*}	1
C	C→A：{*e*,*f*,*g*}，C→B：{*a*}，C→D：{*a*,*g*}	C→A,B,D：{*a*,*e*,*f*,*g*}	2
D	D→B：{*φ*}，D→C：{*c*,*d*}	D→B,C：{*c*,*d*}	0
E	E→F：{*h*}	E→F：{*h*}	0
F	F→E：{*d*,*e*,*f*}	F→E：{*d*,*e*,*f*}	0

2. 网络模型

如图 6-9 所示，在对 NBT 广播数据传输机制的研究中，假设 DTMSN 网络初始部署时，M 个传感器节点随机分布在一个 $l \times l$ 的正方形区域 A 内（虚线代表节点通信半径），基站 BS 位于区域中心，静止不动。节点通信半径为 r，所有节点的移动都遵循随机游走（RWP）移动模型。同时假设节点间的通信信道并非完全理想，数据传输过程中会出现丢包或错误。

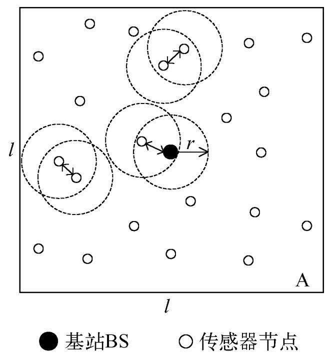
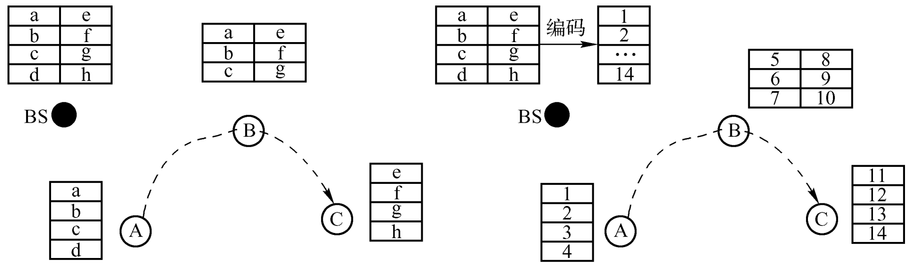


图 6-9　DTMSN 网络模型

3. NBT 基本思想

非编码广播机制如图 6-10（a）所示，假设 BS 有 a～h 共 8 个数据需要广播，节点 A、B、C 在各自运动过程中由于信道等原因仅从 BS 分别获得了 4、6、4 个广播数据，即节点 A 携有数据 a,b,c,d，节点 B 携有数据 a,b,c,e,f,g，节点 C 携有数据 e,f,g,h。由于网络中节点是运动的，在节点 A 沿图中虚线继续运动过程中，当 A 与 B 发生连接时，节点 A、B 都无法从对方处获得全部自身需要广播的数据包，即 A、B 发生连接并交换广播数据后，A、B 节点都因缺少数据 h 而无法获得全部广播数据；只有在节点 A 与 C 发生连接并交换数据后，A、C 才能将广播数据获取完整，而节点 B 则没能获得全部广播数据，需要进一步等待。因此，在 A 沿虚线运动这一时段内，仅有 A、C 获取了全部广播数据。

采用编码广播机制时，如图 6-10（b）所示，BS 利用随机选取的不相关的编码向量将 8 个广播数据编码后发送。同样，假设 A、B、C 在各自运动过程中分别获得了 4、6、4 个广播数据，即节点 A 携有编码数据 1,2,3,4，节点 B 携有编码数据 5,6,7,8,9,10，节点 C 携有编码数据 11,12,13,14。在节点 A 沿虚线继续运动过程中，当 A 与 B 发生连接时，节点 A、B 都能够从对方获取到自身没有的编码数据并以非常高的成功率进行译码，从而获得全部广播数据。在 A 与 C 发生连接后，同样 C 节点也能够成功译码。因此，在 A 沿虚线运动这一时段内，A、B、C 都能够以较高的概率获取全部广播数据，相比非编码机制降低了广播的延迟。



（a）非编码广播　　（b）编码广播

图 6-10　非编码广播和编码广播机制比较图

4. 原始广播数据编码

BS 在广播原始数据前，首先需要根据原始数据的大小将其分为若干个批次，每个批次包含长度相等的 l 个原始数据包。当有传感器节点移动到 BS 通信范围内时，BS 针对每一批次的原始数据包随机选择编码向量对其进行编码，并将编码后的编码包以单播形式发送给传感器节点。BS 对同一批次的数据包每进行一次编码都需要将编码包的序号加 1。需要指出的是，为了提高节点编码包的不相关度，任意两个不同序号的编码包使用的编码向量是不相关的，BS 可以将选择的编码向量与历史编码向量相比较来实现这一目的。同样，BS 以单播形式对编码数据进行传输也是为了降低节点间编码包的不相关度。广播包的格式及所在层次如图 6-11 所示，可以看出广播包的包头处于 MAC 层之上，其中 PACK_TYPE 代表报文类型（这里为广播数据包），B_ID 为本次广播的标识，BATCH_NUM 代表本次广播包含批次的数量，BATCH_NO 为广播数据的批次号，CPACK_NO 是编码包的序号，CODE_VECTOR 为编码向量。

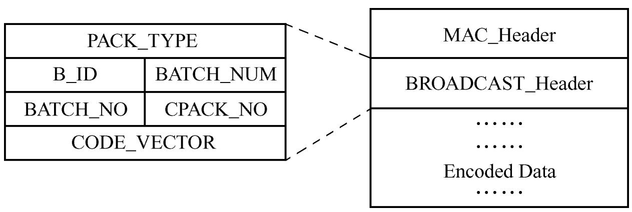


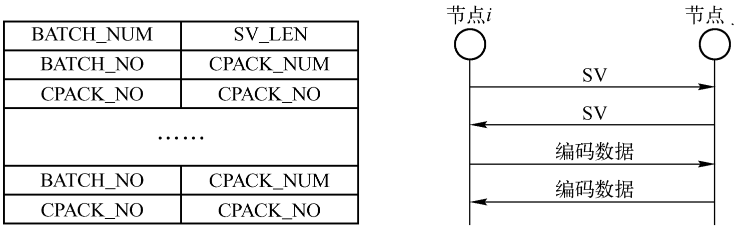
图 6-11　编码广播数据包格式

5. 广播数据交互

为了降低广播延迟，传感器节点之间在彼此进行编码广播数据传输时采用泛洪的机制。当任意两个传感器节点进入彼此通信范围内时，节点间首先通过交换各自的消息索引向量 SV（Summary Vector）来确定自身能够提供给对方的编码数据包及其数量，然后将编码数据包发送给对方（见图 6-12）。假设 a 为节点 i 某批次所缺少编码包的数量，而 b 为邻居节点 j 能够提供的该批次不同序号编码包的数量，则节点 j 向 i 传输编码包的数量为 $\min(a,b)$。实际传输中由于受到信道不确定性的影响，对丢失的广播数据包需要进行重传，NBT 采用简单自动请求重传机制，最高重传次数设定为 m。

向量 SV 格式及任意两节点间的通信过程如图 6-12（a）和图 6-12（b）所示。其中

BATCH_NUM 表示节点自身具有批次的数量，SV_LEN 给出了该 SV 的长度，BATCH_NO 是批次的编号，其后的 CPACK_NUM 表示节点具有该批次编码消息的数量，而 CPACK_NO 则给出节点具体有该批次的哪些编码包。当传感器节点与基站 BS 相遇时，BS 也需要根据节点提供的 SV 来进行广播数据的传输。



（a）SV格式　（b）通信过程

图 6-12　索引向量 SV 格式及通信过程

6. 节点译码

由于节点的移动性，DTMSN 中节点间保持连接的时间通常较短，为了在有限的时间内尽量多地进行数据传输，在 NBT 机制中节点并不是每接收到一个新的编码数据包后就试图解码，而是在获取足够多的编码数据包后才进行解码操作，具体为：节点接收到一个新的编码数据包后，首先判断该批次中自身已有的编码包数量是否达到 l，如果编码包数量达到 l，则提取 l 个编码向量并利用高斯消去法判断编码矩阵是否满秩，满秩则进行解码；否则将冗余向量对应的编码数据包删除，重新向对方请求序号为其他值的本批次编码包。

由于 NBT 中序号不同的编码包采用的编码向量是不同的，因此节点只要接收到 l 个序号不同的编码包就能以很高的概率进行译码，这种对编码包添加序号的方法能够大大节省传感器节点的计算开销和时间，从而保证在较短的连接时间内尽量完成数据传输。

虽然 NBT 机制有着较好的广播时延性能，但也存在广播开销优化策略简单、粗糙的缺点。

参考文献

[1] S. Katti et al., XORs in the Air: Practical Wireless Network Coding, Proc. ACM SIGMOBILE Mobile Comp. Commun. Rev., Oct. 2006: 1－2.

[2] JAGGI S, LANGBERG M, KATTI S, et al. Resilient network coding in the presence of Byzantine adversaries [C] // Proceedings of 27th IEEE International Conference on Computer Communications (INFOCOM'07), Mar 6－12, 2007, Anchorage, AK, USA., Piscataway, NJ, USA: IEEE, 2007: 616－624.

[3] CAI N, YEUNG R W. Network error correction, part II: Lower bounds [J]. Communications in Information and Systems, 2006, 6 (1): 37－54.

[4] YANG S, YEUNG R W. Characterizations of network error correction/detection and erasure correction [C] // Proceedings of Third Workshop on Network Coding, Theory, and Applications (NETCOD'07), Jan 2007, San Diego, CA, USA. 2007.

[5] SILVA D, KSCHISCHANG F R, KOETTER R. A rank－metric approach to error control in random network coding [J]. IEEE Transactions on Information Theory, 2008, 54 (9): 3951－3967.

[6] YEUNG R W, CAI N. Network error correction, part I: Basic concepts and upper bounds [J]. Communications in Information and Systems, 2006, 6 (1): 19 –36.

[7] C. Gkantsidis and P. R. Rodriguez. Network Coding for Large Scale Content Distribution [C]. IEEE INFOCOM'05 (2005).

[8] Cheng L, Das S K, Di Francesco M, et al. Scalable and Energy – Efficient Broadcasting In Multi – Hop Cluster – Based Wireless Sensor Network [C]. The 2011 IEEE International Conference on Communications (ICC 2011), Kyoto, Japan, 2011: 1 –5.

[9] Huang T, Lin Y, Tang L. Neighbor – Aware Gossip – Based Broadcasting Scheme for Wireless Sensor Networks [C]. 2010 International Conference on Communications and Mobile Computing, Shenzhen, China, 2010: 293 –297.

[10] Montolio – Aranda P, García – Alfaro J, Megías D. Improved Flooding of Broadcast Messages Using Extended Multipoint Relaying [J]. Network and Computer Applications, 34 (2), 2011: 542 –550.

[11] Karlsson Gunnar, Lenders Vincent, May Martin, Delay – Tolerant Broadcasting [J]. IEEE Transactions on Broadcasting, 3 (1), 2007: 369 –381.

[12] Fragouli C, Widmer J, and Boudec J L. Efficient Broadcasting Using Network Coding [J]. Presented at IEEE/ACM Transactions on Networking, 16 (2), 2008: 450 –463.

[13] Widmer J, Boudec J L, Network Coding for Efficient Communication in Extreme Networks [C]. In: Proceedings of the ACM SIGCOMM workshop on Delay – tolerant Networking, Philadelphia, PA, USA , 2005: 284 –291.

[14] Pasztor B, Musolesi M, Masxolo C, Opportunistic Mobile Sensor Data Collection with SCAR [C]. In: Proc of the 4th IEEE International Conference on Mobile Ad hoc and Sensor Systems. Piscataway: IEEE Press, 2007: 1 –12.

[15] Goundan A, Coe E, Raghavendra C S, Efficient Broadcasting in Delay Tolerant Networks [C]. In: Proc. GLOBECOM, New Orleans, Louisiana , USA, 2008: 523 –527.

[16] Ahlswede R, Cai N, Li S Y R, et al. Network Information Flow [J]. IEEE Transactions on Information Theory, 46 (4), 2000: 1204 –1216.

[17] Ho T, Medard M, Koetter R, et al. A Random Linear Network Coding Approach To Multicast [J] . IEEE Transactions on Information Theory, 52 (10) , 2006: 4413 –4430.

[18] Katti S, Rahul H, Hu W, et al. Xors in The Air: Practical Wireless Network Coding [J] . IEEE/ ACM Transactions on Networking, 16 (3), 2008: 497 –510.

[19] Nguyen D, Tran T, Nguyen T, et al. Wireless Broadcast Using Network Coding [J] . IEEE Transactions on Vehicular Technology, 58 (2), 2009: 914 –925.

[20] 卢冀，肖嵩，吴成柯．基于机会式网络编码的低时延广播传输算法 [J]. 电子学报, 39 (5) , 2011, pp. 1214 –1219.

[21] 程宏兵，王江涛，杨庚．SPINS 安全框架协议研究．计算机科学，2006.

[22] Erdal Cayirci, C. Rong, 无线自组织网络和传感器网络安全．北京：机械工业出版社，2011.

[23] 杨义先．网络编码理论与技术．北京：国防工业出版社，2009.

[24] 刘云浩，等．物联网导论．北京：科学出版社，2010.

[25] 孙立民，李建中，陈渝，等．无线传感器网络．北京：清华大学出版社，2005.

[26] 杨义先，郭钦．网络编码在网络安全中的应用 [J]. 中兴通信技术，2009.

第 7 章　感知层 MAC 协议安全

本章重点分析了物联网中无线传感器网络 MAC 层 802.15.4 协议和无线局域网的 802.11 协议的安全性。本章对 802.15.4 中的超帧机制进行了深入分析，指出了信标广播机制和保证时隙机制中普通节点与协调器节点之间存在缺少身份认证这一安全问题；同时归纳总结了 802.11 协议中 WEP、WPA/WPA－PSK 认证机制存在的安全问题。

7.1　无线传感器网络 802.15.4 MAC 层协议

7.1.1　IEEE 802.15.4 标准

1998 年 3 月，IEEE 802.15 协议工作组正式批准成立。在 IEEE 802.15 协议工作组内有四个任务组（Task Group，TG），分别制定适合不同应用的标准。

① 任务组 TG1：制定了 IEEE 802.15.1 协议，在实际应用中被发展为蓝牙（Bluetooth）技术。IEEE 802.15.1 是一个小范围通信、低速率数据传输的网络标准，常用于个人手持终端设备的短距离通信。

② 任务组 TG2：制定了 IEEE 802.15.2 协议，研究 IEEE 802.15.1 协议标准与无线局域网标准（IEEE 802.11）的共存问题。

③ 任务组 TG3：制定了 IEEE 802.15.3 协议，主要研究怎样实现无线个域网中高速率数据传输技术。该标准主要用于高质量影音数据传输，如个人多媒体视频传输、高品质影音输送等。

④ 任务组 TG4：制定了 IEEE 802.15.4 协议，针对低速无线个人区域网络（Low－Rate Wireless Personal Area Network，LR－WPAN）制定标准。该标准以低能耗、低经济开销和低复杂度数据传输为目标，用于规范小范围如个人或家庭内的不同设备间的低速互连。

低速无线个人区域网与无线传感器网络有很多相似的特征，因此国内外许多研究者把 IEEE 802.15.4 协议作为无线传感器网络 MAC 层通信标准进行研究。IEEE 802.15.4 协议为无线传感器网络物理层（PHY）和介质访问控制层（MAC）制定了统一标准，其特点主要如下。

① 规定了三种不同载波频率的数据传输速率，分别为 20kbps、40kbps 和 250kbps。

② 具有星形网络和点对点网络两种拓扑结构，可在这两种基本结构上构建更为复杂的网络。

③ 地址格式分为两种，分别是 16 位和 64 位，其中 64 位地址是全球和唯一的扩展地址。

④ 支持冲突避免的载波多路侦听技术（Carrier Sense Multiple Access with Collision Avoidance，CSMA/CA）。

⑤ 支持确认（ACK）传输机制，保证了通信的可靠性。

7.1.2　IEEE 802.15.4 网络协议栈

可根据开放系统互连模型对 IEEE 802.15.4 网络协议栈进行分层，如图 7-1 所示，每一低层分别实现一部分数据传输功能，并向更高一层的协议提供服务接口。

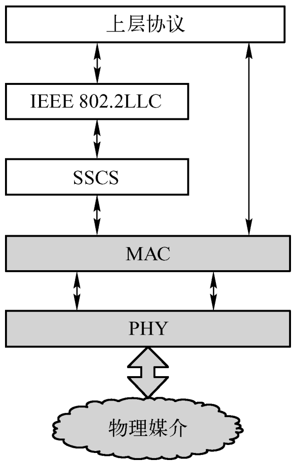


图 7-1　IEEE 802.15.4 协议层图

IEEE 802.15.4 标准为无线传感器网络定义了物理层和数据链路层的介质访问控制子层。其中，物理层由 RF 收发器和一些基础的物理控制部件组成。MAC 子层负责点对点通信接口的管理，主要为高层访问物理信道提供服务。

物理层主要用于提供该层的数据服务和管理服务，它为物理无线信道和 MAC 层之间定义了接口。数据服务主要指的是收发来自无线信道上的数据，管理服务则指的是维护一个由物理层相关数据组成的数据库。物理层数据服务主要有以下五种功能：

① 激活和休眠射频收发器；

② 信道能量检测（energy detect）；

③ 检测接收数据包的链路质量指示（Link Quality Indication，LQI）；

④ 空闲信道评估（Clear Channel Assessment，CCA）；

⑤ 收发数据。

介质访问控制子层主要提供两种服务：MAC 层的数据服务和管理服务。数据服务保证了 MAC 层与物理层之间数据传送的可靠性，管理服务则维护一个存储 MAC 子层协议状态相关信息的数据库。MAC 子层的主要功能有：

① 协调器负责信标帧的构建与广播，其他节点根据接收到的信标帧与协调器节点进行时间同步；

② 支持网络的关联（association）和取消关联（disassociation）操作；

③ 支持无线信道通信安全；

④ 使用 CSMA/CA 机制访问信道；

⑤ 支持保证时隙（Guaranteed Time Slot，GTS）机制；

⑥ 支持不同设备的 MAC 层间可靠传输。

其中，关联是指一个新的设备在加入某一特定无线传感网络时，向协调器提交申请并进行身份认证的过程。无线传感器网络中的节点在从一个网络断开并连接到另一个网络时，就体现了关联和取消关联操作的作用。

7.1.3　IEEE 802.15.4 MAC 帧格式

MAC 层帧结构的设计目标是在多噪声干扰的无线信道环境中，以最低的复杂度实现正确可信的数据通信。每个 MAC 子层的帧结构都包括三部分：帧头（MAC Header，MHR）、负载和帧尾（MAC Footer，MFR）。MAC 帧格式如图 7-2 所示。

字节数：2	1	0/2	0/2/8	0/2	0/2/8	可变	2
帧控制信息	帧序列号	目的设备 PAN 标识符	目标地址	源设备 PAN 标识符	源设备地址	帧数据单元	FCS 校验码
		地址信息					
帧头						MAC负载	MFR帧尾

图 7–2　MAC 帧格式

帧头由帧控制信息（Frame Control）、帧序列号（Sequence Number）和地址信息（Addressing Fields）组成。MAC 子层负载长度是可以变化的，MAC 类型决定了负载的内容。帧头和负载数据的 16 位 CRC 校验码组成了 MAC 子层的帧尾。

MAC 子层设备地址可分为两种格式：一种格式为 16 位（2 字节）的短地址，在设备与网络协调器进行关联操作时，由协调器节点统一分配 16 位短地址，只能保证设备在网内是唯一的，需要结合网络标识符才能起到身份标识作用；另一种格式为 64 位（8 字节）的扩展地址，在设备生产出来之后由厂家设定，是设备全球唯一的标识地址。MAC 帧头长度根据设备所用地址类型不同是可以改变的，这是由于两种地址类型的地址信息长度不一样。帧控制字段的内容决定了该数据帧使用哪种地址类型。由于在物理层的帧里具有说明 MAC 帧长度的字段，因此在 MAC 层帧结构中并没有表示帧长度的字节段。根据物理层帧长和 MAC 帧头的长度可计算得出 MAC 负载长度。

在无线传感器网络中共定义了四种不同类型的帧：信标帧、数据帧、确认帧和命令帧。

1. 信标帧

信标帧格式如图 7–3 所示，其负载数据单元包括四个部分：超帧描述字段、GTS 分配字段、待转发数据目标地址（Pending Address）和信标帧负载。

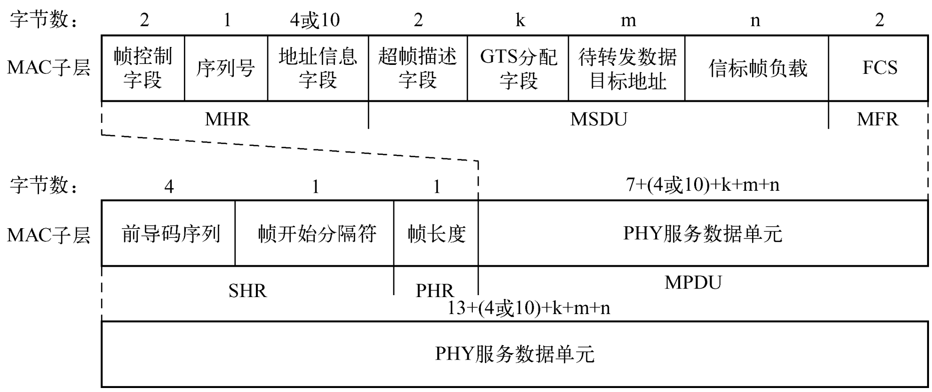


图 7–3　信标帧格式

2. 数据帧

数据帧格式如图 7–4 所示，其负载字段填充了上层需要传送的数据，用来传输上层发到 MAC 子层的数据。数据负载传送至 MAC 子层时，被称为 MAC 服务数据单元（MAC Service Data Unit，MSDU）。在数据帧的首尾分别添上 MHR 头信息和 MFR 尾信息后，就形成了 MAC 帧。

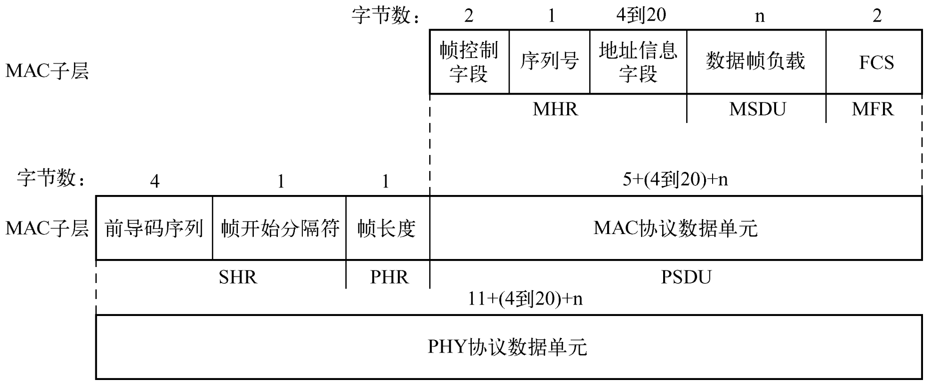


图 7-4　数据帧格式

3. 确认帧

确认帧格式如图 7-5 所示，当设备节点接收到的数据帧或 MAC 命令帧的目的地址为其自身，且数据帧或 MAC 命令帧控制信息字段的确认请求位置设定为 1 时，设备需要回应一个确认帧。IEEE 802. 15. 4 协议规定确认帧的序列号必须与被确认帧序列号一样，同时负载长度必须设置为零。设备在收到被确认帧之后立即返回确认帧，这一操作具有高优先级，不必再利用 CSMA/CA 机制取得信道访问权。

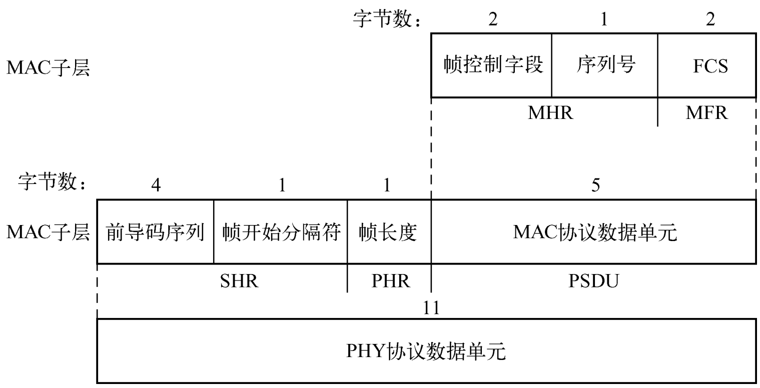


图 7-5　确认帧格式

4. 命令帧

命令帧格式如图 7-6 所示，其功能包括网络构建、数据传输和时间同步等。目前 IEEE 802. 15. 4 协议规定的命令帧类型有九种，主要完成以下三方面的功能：将节点关联传感网络，与协调器节点传递数据，分配保证时隙 GTS。命令帧只是在帧控制字段的帧类型位与其他类型的帧有所不同，在格式上并无大的不同。当帧头的帧控制字段的帧类型为 011b（b 表示二进制数据）时，就表明此帧为命令帧。命令帧的具体功能由帧的负载数据决定。负载数据的结构长度可变，所有命令帧负载的第一字节都是命令类型字节，后面的数据根据命令类型的不同表示不同的含义。

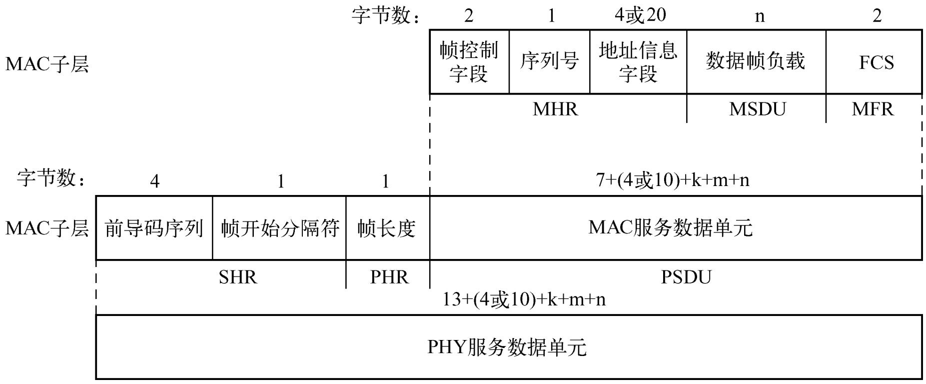


图 7-6　命令帧格式

7.2　IEEE 802.15.4 协议安全分析

7.2.1　信标广播机制及安全分析

在 IEEE 802.15.4 网络中，包括非信标模式和信标模式。在非信标模式下，协调器不再周期性地广播信标，只有当普通节点向协调器发送请求时，协调器才向节点单播信标帧。网络中的普通节点采用非时隙 CSMA/CA 机制竞争接入信道进行数据传输，存在较大的延迟和能耗问题。而在信标模式下，协调器通过周期性地广播信标帧来实现网络节点同步及管理控制，可较好地解决网络通信能耗和时延问题，在实际中得到广泛应用，对其研究比较具有代表性。因此，本书选择信标模式下的 IEEE 802.15.4 网络作为研究对象，下面对网络 MAC 层的信标广播机制、保证时隙（GTS）管理机制进行介绍。

为了支持较强的时效性，满足低延迟或特定数据带宽的通信需求，IEEE 802.15.4 标准提供了一种信标使能模式。在信标模式下，802.15.4 网络配置了一个中心节点（协调器）和一些终端设备节点。其中，协调器利用信标实现下列功能：

① 同步网络节点；

② 管理来自节点的保证时隙（GTS）分配/撤销请求；

③ 为节点传输数据分配专用保证时隙。

终端节点的主要作用是采集传感数据发送给协调器节点。

信标模式下，IEEE 802.15.4 网络使用超帧（Superframe）机制来实现节点的时间同步及数据通信。超帧结构如图 7-7 所示。

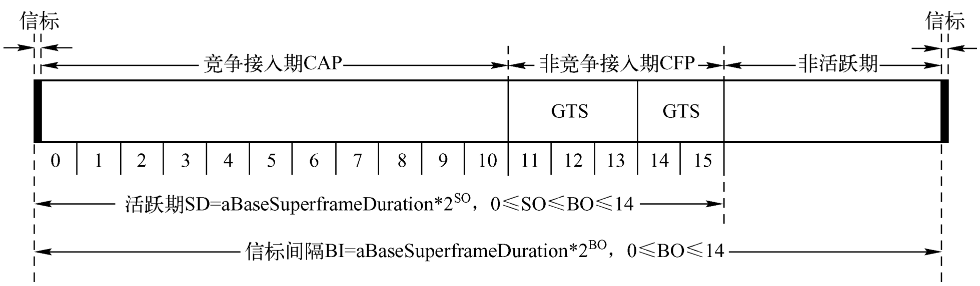


图 7-7　超帧结构

超帧由活跃期和非活跃期两阶段组成，活跃期被划分为 16 个等长时隙（Slot），包括竞争接入期 CAP（Contention Access Period）和非竞争接入期 CFP（Contention Free Period）。在 CAP 期，通信节点使用 CSMA/CA 退避机制接入信道，发送分配/撤销 GTS 请求等；在 CFP 期，节点间无须竞争，各自在分配给它的 GTS 内独享信道进行数据传输，CFP 期最多分配 7 个 GTS；在非活跃期，节点可进入休眠状态以降低能量消耗。

在超帧的第 0 个时隙，网络中的协调器发送一个称为信标（Beacon）帧的广播包，其结构如图 7-8 所示。信标帧包含了超帧持续时长、结构时间划分和 GTS 分配信息，用来实现节点同步和广播 GTS 分配情况。

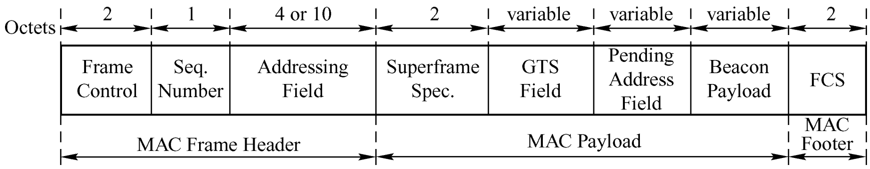


图 7-8　信标帧结构

在信标帧的“Superframe Spec.”域（见图 7-9）中有两个参数 BO（Beacon Order）和 SO（Superframe Order）：信标阶数 BO 决定了信标间隔 BI，即整个超帧的持续时间；超帧阶数 SO 决定了超帧活跃期 SD 的大小。

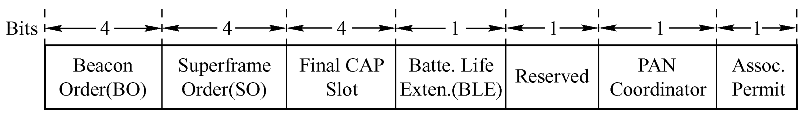


图 7-9　超帧说明域

当网络中的节点接收到信标帧后，提取出 BO 和 SO 值，分别利用下面的公式计算出 BI 和 SD，实现节点与协调器的时间同步。计算公式如下：

$$\text{信标间隔 BI} = \text{aBaseSuperframeDuration} * 2^{BO}$$

$$\text{活跃期 SD} = \text{aBaseSuperframeDuration} * 2^{SO}$$

其中，aBaseSuperframeDuration 为 960symbols，$0 \leqslant SO \leqslant BO \leqslant 14$。

由此可知，“Superframe Spec.”域中的两个参数 BO 和 SO 非常重要，它们决定了信标间隔和活跃期的持续时间长度，在节点与协调器同步过程中起着关键性的作用。然而，当节点收到信标时，仅仅验证信标中的协调器 ID，若和网络初始化时的一致，则认定该信标来自协调器。因此，一旦恶意节点发送带有与协调器相同 ID 的信标帧，合法节点将认定该信标帧同样来自协调器并进行接收。这也暴露了信标广播中存在的最大漏洞：普通节点对协调器的身份不进行认证。

7.2.2　GTS 管理机制及安全分析

信标模式和非信标模式的区别，就在于网络是否提供了一种保证时隙（Guaranteed Time Slot）机制。在保证时隙机制下，IEEE 802.15.4 网络允许节点不必经过 CSMA/CA 竞争信道，在分配给它的 GTS 内可独享信道资源进行数据的传输。这种机制为节点提供了实时性的保障，大大降低了网络时延。

保证时隙 GTS 由协调器负责分配给节点，仅用于被分配 GTS 节点和协调器之间的通信。每个 GTS 可包含一个或多个连续的时隙，协调器最多可同时分配 7 个 GTS。GTS 分配过程是在“先到先服务”（First Come First Serve）模式下进行的，当节点向协调器发送 GTS 分配请求时，协调器根据当前超帧的使用情况来决定是否分配 GTS。当不再使用 GTS 时，可由协调器将其撤销，也可由请求该 GTS 的节点向协调器发送撤销请求来解除使用。在信标帧中，“GTS Field”域有一个 GTS 列表，包含分配到 GTS 的节点设备短地址、GTS 开始时隙及 GTS 长度等信息。普通节点在接收到信标帧后，通过查看 GTS 列表决定何时工作、何时休眠。GTS 分配/撤销过程如图 7-10 所示。

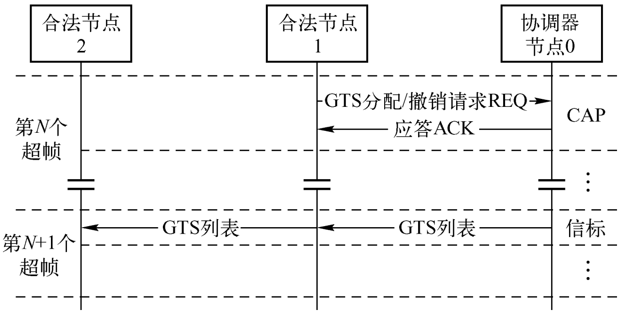


图 7-10　GTS 分配/撤销过程

GTS 分配/撤销请求是在 CAP 期间发送的，普通合法节点 1 首先向协调器节点发送一个 GTS 分配/撤销请求 REQ，协调器在收到请求后向节点 1 发送一个应答 ACK 消息；如果是分配请求，协调器查看当前剩余 GTS 时隙资源，若满足要求则分配 GTS 给节点 1，并将该 GTS 相关参数信息添加到下一信标帧的“GTS Field”域中广播给整个网络。如果是撤销请求，协调器将节点 1 的信息从 GTS 列表中删除，并在下一信标帧中不再包含节点 1 的相关信息。

IEEE 802.15.4 协议规定，协调器负责 CAP 维护，即为满足 CAP 期间有足够时间供节点竞争访问信道，协调器应采用预防措施来保证最小的 CAP 长度 aMinCAPLength。一种措施就是在每 $2n$ 个超帧后撤销未使用的 GTS，其中定义 n 的值为：当 $0 \leqslant \text{macBeaconOrder} \leqslant 8$ 时，$n = 2^{(8-\text{macBeaconOrder})}$；当 $9 \leqslant \text{macBeaconOrder} \leqslant 14$ 时，$n = 1$。然而，这种预防措施存有漏洞，一旦恶意节点不间断地发送 GTS 请求或在 CFP 内利用已分配的 GTS 持续发送数据，协调器将认定该 GTS 一直处于使用状态而不予撤销，该预防措施将因此而失效。

另外，在 CFP 期间，协调器通过管理一个 GTS 列表来控制节点接入网络，负责分配 GTS、撤销 GTS 及避免来自同一合法节点的重复 GTS 请求。协调器管理 GTS 列表主要是管理申请一个或多个 GTS 的节点的 ID，当接收到来自节点的 GTS 申请/撤销请求时，将该节点 ID 添加到 GTS 列表中或从中删除。然而，在处理 GTS 请求时，协调器仅检查传感节点的 ID（一个短 2 字节地址）和报文序列号。因此，若恶意节点伪造一个新 ID 或冒充网络中的合法节点 ID，向协调器发送 GTS 申请/撤销请求，协调器将按正常节点对请求进行处理，必定对合法节点的 GTS 造成影响。这也暴露了协调器在管理 GTS 列表中的漏洞，即缺少对节点 ID 的认证。

7.3　无线局域网概述

无线局域网是无线通信技术与计算机网络技术相结合的产物，是基于无线通信技术在一定的区域范围内将计算机连接起来建立的网络。无线局域网以无线多址信道作为传输媒介，提供与传统有线局域网（Local Area Network，LAN）相同的上网功能，能够使用户真正实现随时、随地、随意地接入宽带网络，是利用无线网络技术实现局域网应用的产物，具有局域网和无线网络两种网络的特点和优势。

7.3.1　无线局域网的基本构成

无线局域网是一个基于蜂窝的架构，每一个蜂窝称为 BSS，其基本构成如图 7-11 所示。每一个蜂窝被一个基站（即访问点或 AP）控制，或是在蜂窝内的点对点网络（Ad Hoc 模式）。

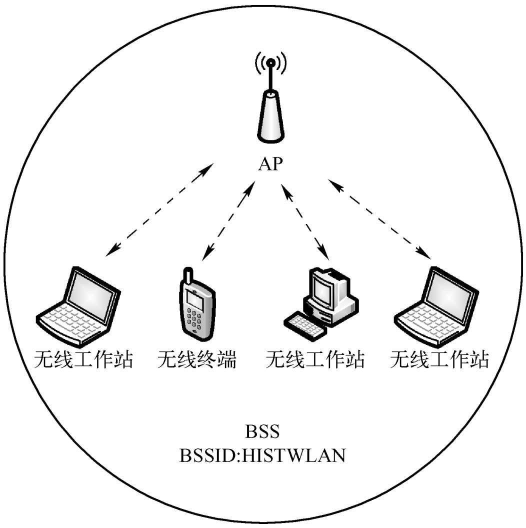


图 7-11　BSS 的基本构成

一个无线分布式系统由多个 BSS 构成，为了标识一个 BSS，通常可以采用给其设置 SSID 来进行区分，这个 SSID 可以称为 BSSID。一个 BSS 可由一个基站和多个无线工作站构成，其中基站指的是无线 AP。通常在无线 AP 上设置 SSID。无线工作站通常指的是安装有无线网络接口的计算机等终端设备。

7.3.2　无线局域网网络结构

无线局域网的网络架构基本上可分为独立型网络结构（Ad Hoc）和基础网络结构两类。

1. Ad Hoc 网络结构

Ad Hoc 网络结构如图 7-12 所示。该结构无须 AP 支持，站点间可相互通信。Ad Hoc 网络结构由一组有无线接口卡的计算机组成。这些计算机以相同的工作组名、ESSID 和密码等

对等的方式相互直接连接，在无线局域网的覆盖范围之内，进行点对点与点对多点之间的通信。

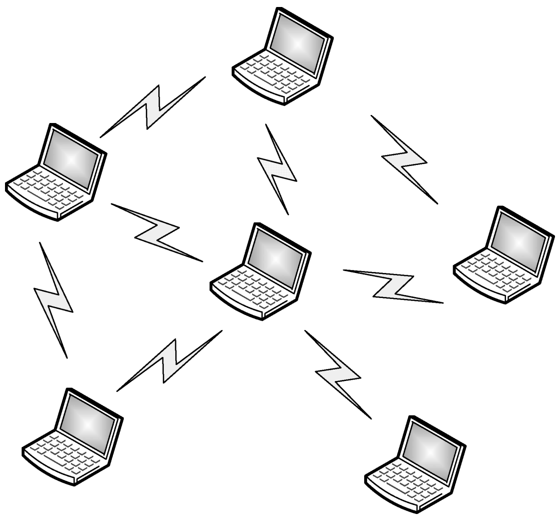


图 7-12　Ad Hoc 网络结构

2. 基础网络结构

基础网络结构指的是具有无线接口卡的无线终端以无线接入点 AP 为中心，通过无线网桥、无线接入网关、无线接入控制器和无线接入服务器等将无线局域网与有线网络连接起来，可以组建多种复杂的无线局域网接入网络，实现无线移动办公的接入。

IEEE 802.11 协议是 IEEE 在 1997 年为无线局域网 MAC 层定义的一个无线网络通信的工业标准。此后这一标准又不断得到补充和完善，形成了 IEEE 802.11x 标准系列。IEEE 802.11x 标准是现在无线局域网的主流标准，也是 Wi-Fi 的技术基础。

7.3.3　IEEE 802.11 相关标准

1. IEEE 802.11

1997 年 6 月，美国电机电子工程师协会（IEEE）为解决无线网络设备互连，制定了 IEEE 802.11 无线局域网标准。IEEE 802.11 是 IEEE 制定的第一个无线局域网标准，主要用于解决办公室和校园网范围内用户终端的无线接入，业务主要限于数据访问。IEEE 802.11 采用 DSSS 或 FHSS 技术，在 RF 射频频段的 2.4GHz 提供了 1Mbps、2Mbps 和许多基础信号传输方式与服务的传输速率规格。由于它在速率和传输距离上都不能满足人们的需要，因此，IEEE 小组又相继推出了其他的 802.11 协议标准。

2. IEEE 802.11b

1999 年 9 月，IEEE 组委会正式批准了 IEEE 802.11b（即 Wi-Fi）标准。此标准规定无线局域网工作频段范围为 2.4～2.4835GHz，数据传输速率达到 11Mbps。IEEE 802.11b 标准

是对 IEEE 802. 11 的一个补充，采用点对点模式和基本网络架构模式两种运作模式，在数据传输速率方面可以根据实际情况在 11Mbps、5. 5Mbps、2Mbps、1Mbps 的不同速率间自动切换，并在 2Mbps、1Mbps 速率时与 IEEE 802. 11 兼容。IEEE 802. 11b 是应用于无线手持终端的最流行的 Wi-Fi 协议，使用直接序列扩频（Direct Sequence Spread Spectrum，DSSS）协议。

3. IEEE 802. 11a

1999 年，IEEE 802. 11a 标准制定完成，此标准规定无线局域网工作频段范围为 5. 15 ~ 5. 825GHz，数据传输速率达到 54Mbps，传输距离控制为 10 ~ 100m。IEEE 802. 11a 与 IEEE 802. 11b 不兼容，采用正交频分复用（OFDM）的独特扩频技术。

4. IEEE 802. 11g

IEEE 802. 11g 是 IEEE 802. 11 工作组于 2003 年推出的无线局域网标准。IEEE 802. 11g 是对 IEEE 802. 11b 的改进。IEEE 802. 11g 工作在 2. 4GHz 频段，采用 OFDM 技术，最大传输速度为 54Mbps，并与 IEEE 802. 11b 兼容。

为防止和 IEEE 802. 11b 设备在无线局域网中共存出现通信冲突的问题，IEEE 802. 11g 协议采用了 RTS/CTS 技术。无线局域网设备在发送数据前，都要先发送一个 RTS（Request To Send）帧，请求使用无线资源，如果这里的 AP 没有和其他设备通信，就发送一个 CTS（Clear To Send）帧，通知它可以与无线 AP 进行通信。

IEEE 802. 11g 与 IEEE 802. 11b 兼容，可共存于同一个 AP 的网络里，从而保障了后向兼容性。这样，原有的无线系统可以平滑地向高速无线网络过渡，延长了 IEEE 802. 11b 设备的使用寿命，减少了用户的开销。

5. IEEE 802. 11n

IEEE 802. 11n 协议采用双频工作模式，工作在 2. 4GHz 和 5GHz 两个频段，这样保障了与以往的 IEEE 802. 11a/b/g 标准兼容。IEEE 802. 11n 标准对 IEEE 802. 11 进行了全面改进，通过采用最新高性能无线技术，优化了网络中数据帧的结构，提高了网络通信的数据吞吐量。

IEEE 802. 11n 理论上的最大速率可达 600Mbps，当前市场上常见的是支持 300Mbps 的 IEEE 802. 11n 的产品。IEEE 802. 11n 标准的核心是 MIMO（Multiple Input Multiple Output，多输入多输出）和 OFDM 技术，使传输速率和传输距离成倍提高。IEEE 802. 11n 网络的最远传输距离可以达到几千米，同时能够保证至少 100Mbps 的传输速率。

表 7-1 列出了 IEEE 802. 11 相关标准的一些参数对比。

表 7-1　IEEE 802. 11 协议标准对比

标准号	802. 11	802. 11b	802. 11a	802. 11g	802. 11n
频率	2. 4 ~ 2. 483GHz	2. 4 ~ 2. 4835GHz	5. 150 ~ 5. 350GHz 5. 475 ~ 5. 725GHz 5. 725 ~ 5. 850GHz	2. 4 ~ 2. 4835GHz	2. 4 ~ 2. 4835GHz 5. 150 ~ 5. 850GHz

续表

标准号	802.11	802.11b	802.11a	802.11g	802.11n
信道数	3	3	24	3	15
速率	2Mbps	11Mbps	54Mbps	54Mbps	600Mbps
频宽	1～2Mbps	20Mbps	20Mbps	20Mbps	20Mbps/40Mbps
调制方式	FHSS/DSSS	CCK/DSSS	OFDM	CCK/DSSS OFDM	MIMO－OFDM DSSS/CCK
兼容性	—	802.11b	802.11a	802.11b/g	802.11a/b/g/n

此外，还有 IEEE 802.11e/f/h/i 等协议标准，分别用于不同方面。

7.3.4　IEEE 802.11 协议体系

如图 7-13 所示，IEEE 802.11 协议主要工作在 OSI 网络模型的物理层和数据链路层。其中，数据链路层由逻辑链路层（Logic Link Control，LLC）和媒体控制层（Media Access Control，MAC）两个子层构成。数据链路层位于物理层上层，基于物理连接建立数据链路，对下可传输具有一定意义和结构的信息，对上可为网络层提供有效服务。数据链路层实现了实体间数据的可靠传输，功能主要包括数据成帧、时间同步、差错控制、流量控制和链路管理。

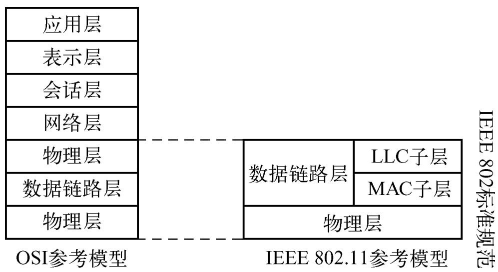


图 7-13　IEEE 802.11 基本结构模型

7.4　无线局域网 MAC 层接入认证协议

7.4.1　WEP 身份认证协议

IEEE 802.11 标准为 WEP 身份认证协议定义了两种方式：开放系统认证（Open System Authentication）和共享密钥认证（Shared Key Authentication）。

1. 开放系统认证

开放系统认证未提供任何安全防护机制，应用于安全级别要求低的场景，整个连接过程分为请求和响应两步。

2. 共享密钥认证

共享密钥认证是一种用于认证的密码技术，它是基于客户端是否具有共享密钥的简单的

“请求/响应” 机制。

如图 7-14 所示，客户端先向接入点 AP 发送一个认证请求，AP 收到这个请求后，通过某种机制产生一个随机数，并将其发送至无线客户端，再结合与 AP 共享的密钥（WEP 密钥)，客户端利用 RC4 流密码算法加密这个随机数，将密文返回给 AP。随后，AP 使用共享密钥对收到的密文进行解密，再将解密结果与发送的随机数相比较，若相同则验证了客户端的合法身份，从而允许其访问网络。否则，拒绝该客户端的访问请求。在加密的过程中，用到的加密算法为 RonRivest 开发的 RC4 流密码算法，该算法不提供双向认证。也就是说，只能进行接入点 AP 对客户端的认证，无法保证通信接入点的合法性。

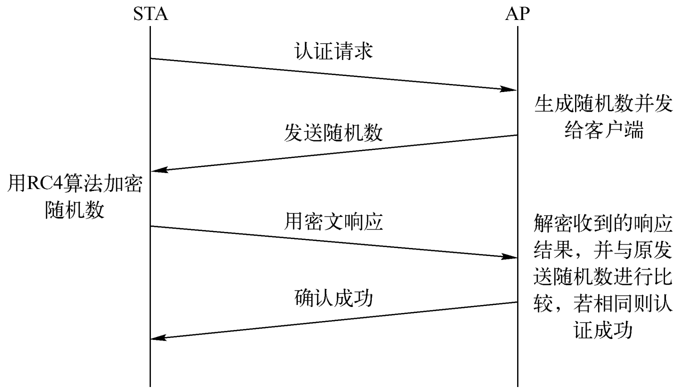


图 7-14　WEP 认证过程

7.4.2　WPA/WPA2 - PSK 认证机制

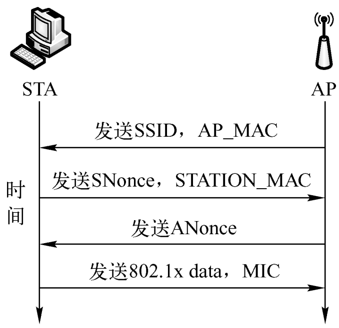


图 7-15　“四次握手” 过程

对于一些中小型的企业网络或家庭用户，WPA/WPA2 提供了一种简化的认证模式，这种认证模式并不需要专门的认证服务器，仅要求预先在每个 WLAN 节点（AP、无线路由器、网卡等）中输入一个密钥即可实现。这种模式称为 WPA/WPA2 预共享密钥（WPA/WPA2 - PSK)，AP 对客户端的身份认证主要通过一个 “四次握手” 过程完成。这种 “四次握手” 过程如图 7-15 所示。

第 1 步，AP 向 STA 发送消息 A，消息 A 仅发送 ANonce 给 STA。消息 A 的内容见表 7-2。

表 7-2　消息 A 的内容

描述符类型	254
密钥信息	请求，错误，安全：0 MIC：0 ACK：1 安全：0 索引：0 密钥类型：成对密钥 描述符类型：1

续表

描述符类型	254
密钥长度	实际密钥长度
重复计数器	< 当前值 >
Key Nonce	ANonce
EAP - Key IV	0
RSC	0
密钥标志符	0
密钥 MIC	0
密钥数据长度	0

第 2 步，STA 得到 ANonce 的复制，并产生它自己的 SNonce。然后计算出临时密钥 PTK。向 AP 发送消息 B，消息 B 仅仅是发送 SNonce 给 AP。消息 B 的内容见表 7-3。

表 7-3　消息 B 的内容

描述符类型	254
密钥信息	请求，错误，安全：0 MIC：1 ACK：0 安全：0 索引：0 密钥类型：成对密钥 描述符类型：1
密钥长度	实际密钥长度
重复计数器	< 当前值 >
Key Nonce	SNonce
EAP - Key IV	0
RSC	0
密钥标志符	0
密钥 MIC	MIC 值
密钥数据长度	密钥数据长度
密钥数据	信息要素

在这步中消息完整性校验 MIC 的值由 HMAC - SHA - 1 算法计算。MIC 的计算包括了从首部的 EAPOL 协议版本字段一直到包含的所有密钥数据。为了计算 MIC，申请者需要使用 PSK，如果 STA 不知道正确的 PSK，它产生的 MIC 不能与期望的 MIC 相符。因此，MIC 在这消息中完成了两件事情：保护消息不受篡改和证实 STA 知道 PSK。

第 3 步，AP 收到消息 B 之后，提取出 SNonce，随后计算临时密钥，之后向 STA 发送消息 C。消息 C 的内容见表 7-4，它主要提供了两个功能：首先，用于 STA 确认 AP 是否知道 PSK；其次，通知 STA，AP 将要准备安装和使用临时密钥。直到 AP 收到第 4 步消息 D 后，它才开始实际安装临时密钥。

表 7-4　消息 C 的内容

描述符类型	254
密钥信息	请求，错误，安全：0 MIC：1 ACK：1 安全：0 索引：0 密钥类型：成对密钥 描述符类型：1
密钥长度	实际密钥长度
重复计数器	<当前值>
Key Nonce	ANonce
EAP－Key IV	0
RSC	开始的序列号
密钥标志符	0
密钥 MIC	MIC 值
密钥数据长度	密钥数据长度
密钥数据	信息要素

第 4 步，STA 向 AP 发出确认准备安装临时密钥的消息 D。消息 D 的内容见表 7-5。

表 7-5　消息 D 的内容

描述符类型	254
密钥信息	请求，错误，安全：0 MIC：1 ACK：0 安全：1 索引：0 密钥类型：成对密钥 描述符类型：1
密钥长度	实际密钥长度
重复计数器	<当前值>
Key Nonce	SNonce
EAP－Key IV	0
RSC	开始的序列号
密钥标志符	0
密钥 MIC	MIC 值
密钥数据长度	密钥数据长度

7.4.3　IEEE 802.1x/EAP 认证机制

对于大型企业的应用，常采用 IEEE 802.1x/EAP 的方式，用户提供认证所需的凭证。基于端口的接入控制协议（Port Based Network Access Control Protocol）即 IEEE 802.1x 协议，源于有线网络，提供了可靠的用户认证和密钥分发框架，可对无线网络的用户进行身份验证

和访问控制。可扩展认证协议 EAP（Extensible Authentication Protocol）由点对点协议（PPP）扩展而来，建立在挑战 - 响应的通信模型上。

1. IEEE 802.1x 基于端口的接入控制协议

IEEE 802.1x 是 IEEE 为有线网络提供的一种基于端口的接入控制协议，同样适用于无线网络。IEEE 802.1x 的认证过程定义了如下参与者。

① 端口接入实体是在端口捆绑的请求认证实体和认证协议实体，负责支持认证双方的认证过程。

② 申请者是申请接入的无线客户端。在要发起 IEEE 802.1x 协议的认证过程时，客户端必须运行 IEEE 802.1x 客户端协议软件并提供必要的认证信息（如身份标识和口令信息等）。

③ 认证者是负责将 STA 与网络分离的设备，以防止非法访问，无线局域网中对应的设备为 AP。它在认证环节负责中继申请者与认证服务器间的交互信息，在认证环节结束后，充当普通的接入点角色。

④ 认证服务器处于整个认证体系结构的核心位置，是负责 STA 认证的后端设备，根据数据库中的信息记录处理接入请求。无线局域网中对应的设备是 RADIUS 服务器。

IEEE 802.1x 协议区别于其他认证协议的特征是其具有“受控端口”和“非受控端口”的概念，访问端口分为物理端口和逻辑端口两种情况。IEEE 802.1x 协议中的端口控制原理结构如图 7-16 所示。

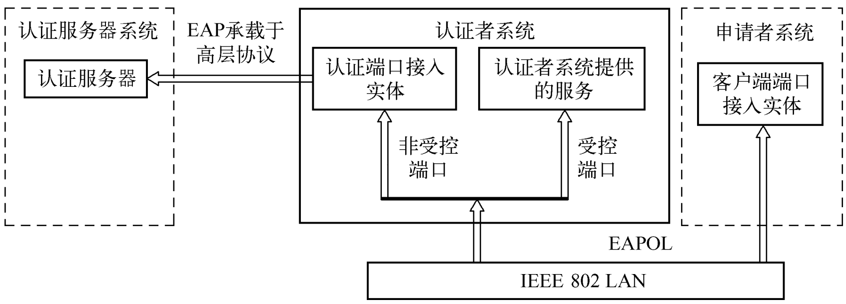


图 7-16　IEEE 802.1x 协议中的端口控制原理结构

2. EAP 可扩展认证协议

IEEE 802.1x 协议的认证是基于可扩展认证协议（Extensible Authentication Protocol, EAP）实现的。EAP 实现了 PPP 认证时间的延迟，利于在决定采取何种机制前请求到更为丰富的信息。相对于定性为认证协议，其实可以说 EAP 是一种认证协议的封装格式，通过 EAP 客户端和认证服务器间能够实现对具体认证协议的动态协商。

EAP 包含 4 种基本报文，即 EAP Request、EAP Response、EAP Success 和 EAP Failure，在局域网中对应的是 EAPOL（EAP over LAN）。IEEE 802.1x 消息利用两种 EAP 方式传输：一是在 STA 与 AP 间的链路上运行 EAPOL 协议；二是在 AP 与 AS 间采用封装到高层协议的 EAP 协议。IEEE 并没有定义它自己的协议，现在大部分都采用 EAP over RADIUS（Remote Authentication Dial In User Service）标准。

图 7-17 所示为一个典型的 IEEE 802.1x/EAP 中的协议实体。IEEE 802.1x/EAP 将 EAP

报文承载在 AP 与 RADIUS 服务器间通信所使用的 RADIUS 协议中，该协议提供了对每个数据包层面的认证和完整性校验机制。

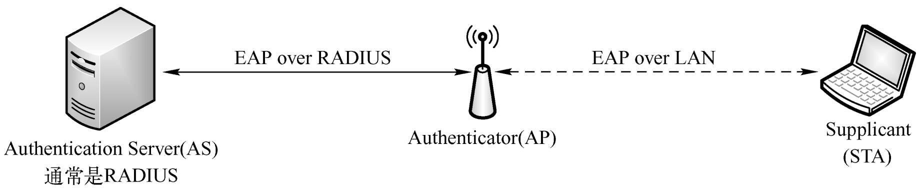


图 7-17　IEEE 802. 1x/EAP 中的协议实体

RADIUS 协议在 IEEE 802. 1x 中的认证架构如图 7-18 所示。认证者收到来自未经过认证的申请者接入请求，只允许申请者的认证信息包通过而拒绝申请者接入网络，将接入请求信息交付给后台的认证服务器进行操作。

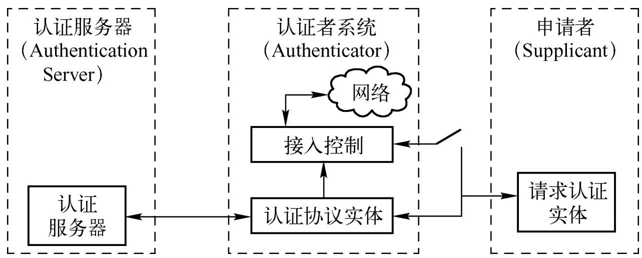


图 7-18　RADIUS 协议的认证架构

申请者通过 EAP 协议与认证服务器进行认证交互，认证数据封装在 EAP 协议包，其结构如图 7-19 所示。认证者负责将 EAP 协议封装到如 RADIUS 等其他高层协议中，提高 EAP 协议到达认证服务器的成功率。

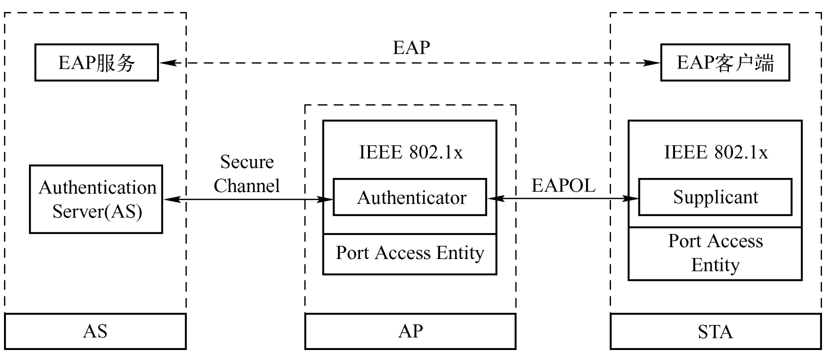


图 7-19　IEEE 802. 1x 认证结构

WLAN 下 IEEE 802. 1x 协议的双向认证流程如图 7-20 所示。

在 IEEE 802. 1x/EAP 认证过程中，STA 首先发送 EAPOL - start 给 AP，成功接收到消息后，AP 发送 EAP - Request/Identity 给 STA，要求 AP 发送其身份。STA 返回 EAP - Response/Identity 作为身份请求消息的应答。AP 将应答信息中继给 AS。接下来 STA 和 AS 间启动认证消息的交互。认证过程结束后，AS 决定允许还是拒绝 STA 的访问，AS 通过 EAP -

Success 或 EAP－Failure 来通知 STA 最后的结果。在 AP 转发 Success/Failure 消息时，它也根据此消息来允许或阻止 STA 通过它的数据流。如果认证成功，STA 和 AS 会得到一个主密钥 MK（Master Key），同时 STA 和 AP 会得到一个共享密钥 PMK（Pairwise Master Key）。

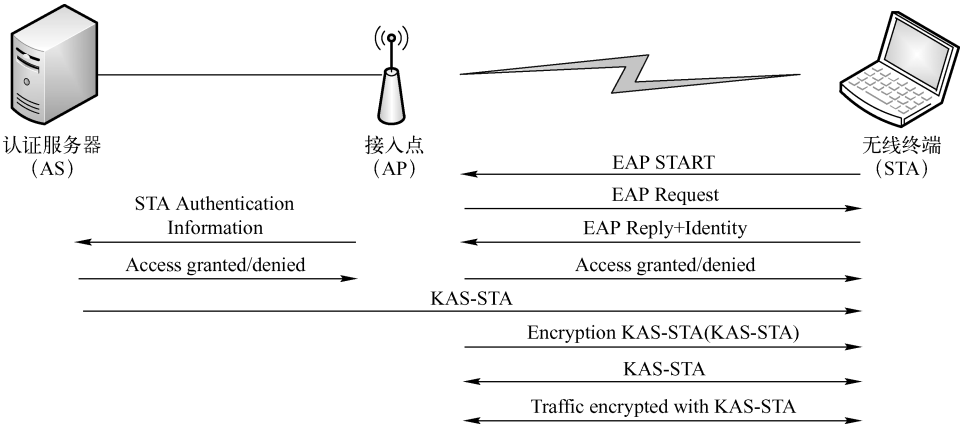


图 7-20　WLAN 下 IEEE 802. 1x 协议的双向认证流程

出于对传输层安全的考虑，EAP 允许采用灵活的方案，并且对认证技术向后兼容，这主要是通过扩展“EAP 类型”域实现的。简言之，EAP 协议是一个协议载体，支持以其为平台而开发更为合适的认证方法。IEEE 802. 1x 中 EAP 协议应用的层次结构如图 7-21 所示。

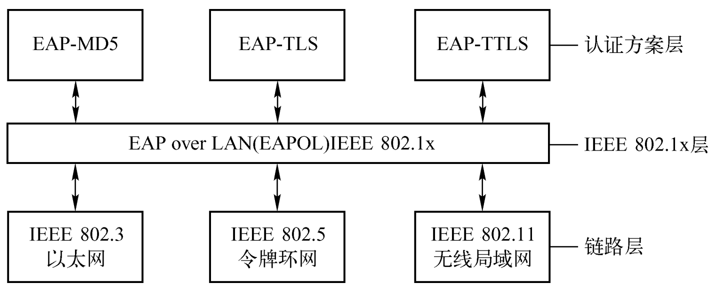


图 7-21　802. 1x 中 EAP 协议应用的层次结构

目前最常用的 EAP 方式有 EAP－MD5、EAP－TLS、EAP－TTLS、EAP－PEAP 和 EAP－SIM 等。

（1）EAP－MD5（消息摘要认证协议）

EAP－MD5 是 EAP 支持的一种传统的认证协议。该协议采用单向认证机制，只支持客户端到认证服务器的认证。协议认证中，服务器只核查客户端的用户名及口令，若和数据库的信息吻合则通知认证系统 AP 放行该客户端，否则不允许客户端接入网络。

（2）EAP－TLS（传输层安全认证协议）

EAP－TLS 是基于 TLS 的认证协议，与 EAP－MD5 所不同的是其支持双向认证。协商会话密钥阶段采用 TLS 协议。协议以公钥证书为基础进行 TLS 建立会话密钥。EAP－TLS 是 IEEE 802. 11i 协议中支持的默认认证方法，具有较高的安全性，配置比较复杂。

（3）EAP－TTLS（隧道传输层安全认证协议）

EAP－TTLS 由 Funk Software 和 Certicom 公司开发，扩展了 EAP－TLS 认证协议。此协

议支持基于证书的认证机制，与 EAP - TLS 的区别是该协议只需要认证服务器的证书。

（4） EAP - PEAP（受保护的可扩展认证协议）

EAP - PEAP 认证过程中只需要认证服务器端的证书，精简了安全体系结构。PEAP 只对认证服务器具备公共密钥有要求，认证过程中交换以 TLS 报文格式封装的 EAP 报文，报文被协商的 TLS 会话密钥进行保护和确认。

（5） EAP - SIM（基于用户身份识别模块的认证协议）

EAP - SIM 是一种基于 SIM 卡的认证机制，应用广泛。EAP - SIM 最大的特点是通过手机 SIM 卡信息实现身份认证，用户只需要通过增强的 GSM 认证算法验证即可接入网络。

IEEE 802.1x/EAP 常用认证协议比较见表 7-6。

表 7-6　802.1x/EAP 常用认证协议比较

认证协议	MD5	TLS	TTLS	PEAP	SIM
服务器认证	否	证书	证书	证书	GPRS/GSM
客户端认证	口令	证书	口令或证书	证书（可选）	Sim 卡
双向认证	否	是	是	是	是
认证方式	挑战/响应认证	TLS 会话双向认证	认证服务器使用证书，客户端在加密隧道中完成认证	类似于 TTLS，基于 MSchapv2 认证协议	使用 Sim 卡
密钥管理	否	是	是	是	否
配置难度	易	难	一般	一般	易
安全性	低	高	一般	一般	高

3. RADIUS 协议

RADIUS 即远程拨号接入用户认证协议（Remote Authentication Dial In User Service），是以客户 - 服务器为模式的安全协议。实现将用户信息中继给 RADIUS 服务器的功能。

为了能够支持 EAP，RADIUS 在原有协议的基础上增加了两个新的属性：EAP 消息（EAP - Message） 和认证消息（Message - Authenticator）。许多认证方案被添加进来，主要是通过新的属性 EAP - Message 来实现的。EAP - Message 允许在不调整它的结构和改动 EAP 消息的前提下，把 EAP 消息封装进 RADIUS 报文，RADIUS 服务器将会把 EAP 消息发送给后端的认证服务器，两个服务器也可以是同一台机器，它们用适当的协议通信。

客户端和网络接入服务器 NAS（WLAN 中指的是 AP）之间的 EAP 会话以 LCP（链路控制协议）协商 EAP 开始，通常起始时发送 EAP 身份请求，具体实施步骤如下。

Step1：NAS 向认证端点发送一个 EAP 身份请求（Request/Identity） 包。

Step2：客户端回应一个 EAP 身份响应（Response/Identity） 包。

Step 3：NAS 得到用户信息后，会根据 RADIUS 标准规定的格式，向 RADIUS 服务器发送一个接入请求（Access Request） 包。若 RADIUS 设置为代理服务器，RADIUS 服务器还可将数据包转向其他的认证服务器。

Step 4：如果 RADIUS 服务器支持 EAP，则回应一个质询或接受数据包；否则回应拒绝访问。其中的 EAP 消息属性将会继续传送给客户端。如果 Access - Request 包含了用户名，

为了允许通过 NAS 传送回复访问，RADIUS 服务器的并发接受访问数据包中也必须包括用户名，否则很难进行计费和次序的处理。

Step 5：当客户端接收到 RADIUS 的拒绝或接受数据包，会话结束。不管其中是否包含 EAP 认证成功或失败消息，NAS 都必须向端点发送一个 LCP 结束请求。RADIUS 接受数据包表明 EAP 认证成功，必须包含所有期望的属性。

EAP/RADIUS 工作流程如图 7-22 所示。

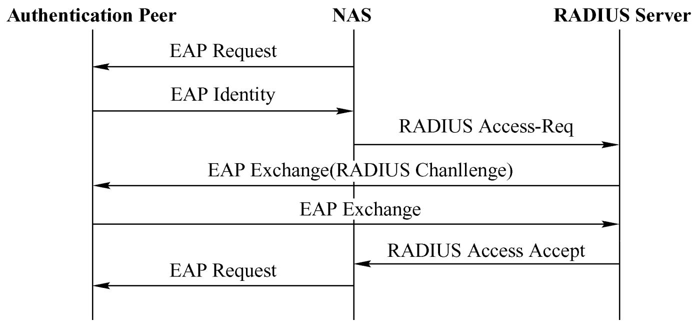


图 7-22　EAP/RADIUS 工作流程

7.5　无线局域网 MAC 层协议安全分析

7.5.1　WEP 中的安全隐患

WEP 中存在安全隐患已被相当多的研究者发现提出，可以概括为：

① RC4 算法的缺陷主要有密钥重复、弱密钥问题和密钥生成简单等问题；

② IV 空间问题主要是 IV 空间最多只有 2^{24} 个，即只有 16 777 216 个值，容易引起 IV 冲突，造成密钥泄露。

在 WEP 加密协议保护下，当客户端尝试连接 AP 时，AP 首先向客户端发送一个 Challenge packet，客户端使用共享密钥将此值加密后再返还给 AP，AP 对收到的数据和其自运算得到的值进行比较，只有两个数据值相同，才允许客户端接入无线网络。攻击者只要收集足够的数据包，从中提取出 IV 值和密文，而与这个密文对应的明文的第一字节是确定的，为逻辑链路控制的 802.2 头信息，通过这一字节的明文和密文，做异或运算就能得到一字节的 WEP 密钥流。由于 RC4 流密码产生算法只是把原来的密码打乱次序，所以获得的这一字节的密码就是 IV + PASSWORD 的一部分。虽然打乱了 RC4，不知道这一字节具体的位置和排列次序，但当攻击者收集到足够多的 IV 值还有碎片密码时，就可以进行统计分析，用上面的密码碎片重新排序配合 IV 使用 RC4 算法得出的值和多个流密码位置进行比较，最后得到这些密码碎片正确的排列次序。至此，WEP 的密码就被分析破解出来了。

7.5.2　WPA/WPA - PSK 认证协议安全分析

从“四次握手”过程可以看出，在 WPA 中与密码有关联的是四次握手包，在四次握手中主要传递的有如下数据：SSID，AP_MAC，STATION_MAC，SNonce，ANonce，802.1xdata，

MIC。很明显，前六个元素一般不涉及密码的内容，只有经过多次运算得到的 MIC 值携带密码的相关信息。

图 7-23 为 MIC 派生图，表明了 MIC 值的具体运算过程，容易得知 MIC 值是利用 SSID，AP_MAC，STATION_MAC，SNonce，ANonce，802.1x data 和密码基于三个算法派生得到的。

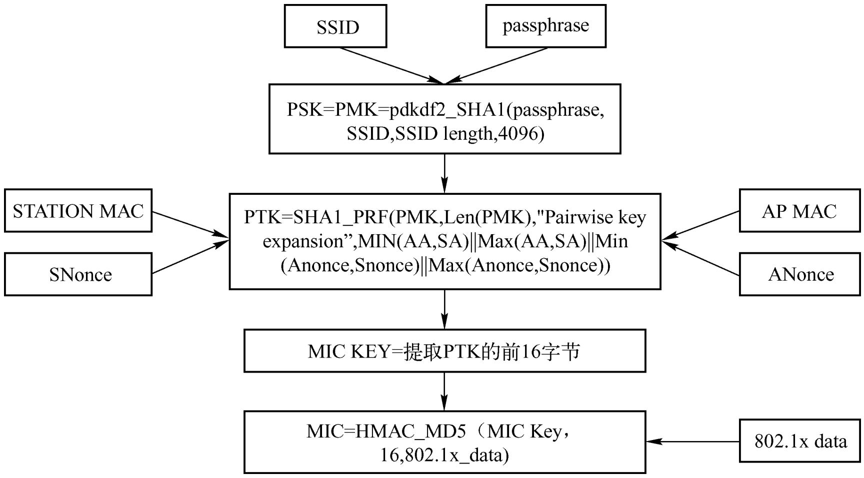


图 7-23　MIC 派生图

MIC 是在 WPA－PSK4 次握手过程的第 2 步中基于 PSK 计算得到的，其中 PSK 是有限长度密码。若获得 WPA－PSK 握手数据包，即获得 MIC 的值，则可以采用“字典攻击”法来实现对 PSK 密钥的破解。破解时间取决于 PSK 密钥长度，一般一个 5 位数 PSK 密钥的“字典攻击”在主流单机环境下所花费的时间为 4～23 天。

“字典攻击”的原理为：对猜测的 PSK，通过 PBKDF2 函数可计算出 PMK，再调用 4 次 HMAC－SHA－1 函数计算得到 PTK，最后调用 1 次 HMAC－SHA－1 函数得到 MIC 的值，用计算出的 MIC 和捕获到的握手信息做比较，如果一致，则表明 PSK 密钥猜测成功，即得到 PSK；若猜测不成功，则继续猜测有可能使用的 PSK。

若采用简单“字典攻击”，则破解 WPA－PSK 密钥非常慢，若 WPA－PSK 密钥再长一些，则“字典攻击”就没有太多的实战能力甚至失去破解的意义。后来的研究者通过对算法的研究改进了“字典攻击”效率。提高效率的途径是预先利用目标 SSID 使用相同的算法生成一个 Hash 字典库，在需要破解时，直接将捕获的相关数据与 Hash 字典库里的值进行比较，有效避免了算法中循环计算所耗费的时间，大大提高了破解效率。这种预运算的破解方式，可以将破解的效率提高 200～3000 倍。

7.5.3　IEEE 802.1x/EAP 认证协议安全分析

IEEE 802.1x 认证机制比之前的认证协议具有更高的安全性，它增加了一个 RAIDUS 认证服务器，可对网络中的用户实行集中式管理。然而，IEEE 802.1x 是为有线网络设计的，未充分考虑无线网络的特点，且只是实现了客户端与认证服务器之间的单向认证，客户端与认证系统 AP 之间缺少安全认证。单向认证基于物理因素在有线环境下或许不存在问题，但在开放的无线网络环境中就可能受到中间人攻击。攻击者通过假冒成 AP 欺骗客户端与其进

行数据通信，也可通过伪造 MAC 地址假扮合法客户端与认证服务器进行通信。因此，IEEE 802.1x/EAP 需要结合其他安全协议才能为网络提供安全可靠的接入认证，比较经典的就是基于 RADIUS 服务器证书认证的 EAP－TLS 认证协议。

参考文献

[1] Akyildiz I F, Su W, Sankarasubramaniam Y, et al. A survey on sensornetworks [J]. Communications magazine, IEEE, 2002, 40 (8): 102－114.

[2] Nicopolitidis P, Pomportsis A S, Papadimitriou G I, et al. Wireless Networks [M]. Chichester: John Wiley & Sons, 2003.

[3] IEEE 802 Working Group. Standard for Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (LR－WPANs) [S]. ANSI/IEEE 802.15, 2003, 4.

[4] Akyildiz I F, Vuran M C. Wireless sensor networks [M]. Chichester: John Wiley & Sons, 2010.

[5] Anastasi G, Conti M, Di Francesco M. A comprehensive analysis of the MAC unreliability problem in IEEE 802.15.4 wireless sensor networks [J]. Industrial Informatics, IEEE Transactions on, 2011, 7 (1): 52－65.

[6] IEEE Standards Board. 802 part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Standard 802.11, 1999, 1999 Edition.

[7] Weatherspoon S. Overview of IEEE 802.11 bSecurity [J]. Intel Technology Journal Q, 2000, 2: 2000.

[8] 王建平，余根坚，李晓颖等．无线网络技术 [M]．北京：清华大学出版社，2013.

[9] IEEE Standard for Information Technology－Telecommunications and Information Exchange between Systems－Local and Metropolitan Srea Networks－Specific Requirements Part 15.1: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs), IEEE Std 802.15.1－2002.

[10] IEEE Recommended Practice for Information Technology－Telecommunications and Information Exchange between Systems－Local and Metropolitan Area Networks－Specific Requirements Part 15.2: Coexistence of Wireless Personal Area Networks with Other Wireless Devices Operating in Unlicensed Frequency Bands, IEEE Std 802.15.2－2003.

[11] IEEE 802.15.3. Wireless MAC and PHY Specifications for High Rate WPAN, IEEE 802.15.3TM, 2003.

[12] IEEE Standard for Information Technology Elecommunications and Information Exchange between Systems－Local and Metropolitan Area Networks－Specific Requirements, Part 15.4: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low－Rate Wireless Personal Area Networks (LR－WPANS), IEEE Std 802.15.4－2003.

[13] Norman Abramson. The ALOHNET－Surfing for Wireless Data [J]. IEEE Communications Magazine, vol.47, no.12, pp.21－25, 2009.

[14] IEEE 802.11 WORKING GROUP (2003) Draft Supplement to Standard for Telecommunications and Information Exchange between Systems－LAN/MAN Specific Requirements－Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications. New York USA: The Institute of Electrical and Electronics Engineers, Inc. 2003.

[15] IEEE 802.11 Working Group. IEEE 802.11－1997: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1997.

[16] IEEE 802. 11 Working Group. IEEE 802. 11a – 1999：High – speed Physical Layer in the 5 GHz band，1999.

[17] IEEE 802. 11 Working Group. IEEE 802. 11b – 1999：Higher Speed Physical Layer Extension in the 2. 4 GHz band，1999.

[18] IEEE 802. 11 Working Group. IEEE 802. 11g – 2003：Further Higher Data Rate Extension in the 2. 4 GHz Band，2003.

[19] Wi-Fi Alliance. Wi-Fi Protected Access security sees strong adoption：Wi-Fi Alliance takes strong position by requiring WPA security for product certification. January 5，2004.

[20] IEEE 802. 11 Working Group. IEEE 802. 11n – 2009：Enhancements for Higher Throughput，2009.

[21] Tsukizawa T，Shirakata N，Morita T，et al. A fully integrated 60GHz CMOS transceiver chipset based on Wi-Gig/IEEE 802. 11ad with built – in self calibration for mobile applications [C] //Solid – State Circuits Conference Digest of Technical Papers（ISSCC），2013 IEEE International. IEEE，2013：230 – 231.

[22] Arbaugh W A. An inductive chosen plaintext attack against WEP/WEP2 [J]. IEEE document，2001，802（01）：230.

[23] Borisov N，Goldberg I，Wagner D. Security of the WEP algorithm [D]. University of California at Berkeley，Feb. 2001.

[24] Fluhrer S，Mantin I，Shamir A. Weaknesses in the key scheduling algorithm of RC4 [C] //Selected areas in cryptography. Springer Berlin Heidelberg，2001：1 – 24.

[25] Arbaugh W A，Shankar N，Wan Y C J，et al. Your 802. 11 wireless network has no clothes [J]. Wireless Communications，IEEE，2002，9（6）：44 – 51.

[26] Alliance W F. Wi-Fi Protected Access：Strong，standards – based，interoperable security for today's Wi-Fi networks [J]. Retrieved March，2003（1）：2004.

[27] Alliance W F. Wi-Fi CERTIFIED™ for WMM™ – Support for Multimedia Applications with Quality of Service in Wi-Fi® Networks [J]. Austin，Wi-Fi Alliance，2004.

[28] Walker J. 802. 11 security series part ii：The temporal key integrity protocol（tkip）[J]. Intel Corporation，2002.

[29] Fluhrer S，Mantin I，Shamir A. Attacks on RC4 and WEP [J]. RSA Laboratories，Cryptobytes，2002，5（2）.

[30] Lashkari A H，Danesh M M S，Samadi B. A survey on wireless security protocols（WEP，WPA and WPA2/802. 11 i）[C] //Computer Science and Information Technology，2009. ICCSIT 2009. 2nd IEEE International Conference on. IEEE，2009：48 – 52.

[31] 中华人民共和国国家标准. 信息技术系统间远程通信和信息交换局域网和城域网特定要求第 11 部分：无线局域网媒体访问控制和物理层规范 [S]. 2003.

[32] 郭渊博，杨奎武，张畅等. 无线局域网安全：设计及实现 [M]. 北京：国防工业出版社. 2010.

[33] Pelechrinis K，Iliofotou M，Krishnamurthy S V. Denial of service attacks in wireless networks：The case of-jammers [J]. Communications Surveys & Tutorials，IEEE，2011，13（2）：245 – 257.

[34] DeBruhl B，Tague P. How to jam without getting caught：Analysis and empirical study of stealthy periodic jamming [C] //Sensor，Mesh and Ad Hoc Communications and Networks（SECON），2013 10th Annual IEEE Communications Society Conference on. IEEE，2013：496 – 504.

[35] Butani B，Shukla P K，Silakari S. Optimized and Executive Survey of Physical Node Capture Attack in Wireless Sensor Network [J]. International Journal of Computer Network and Information Security（IJCNIS），2014，6（11）：26 – 31.

[36] Wilhelm M，Martinovic I，Schmitt J B and Lenders V. Reactive Jamming in Wireless Networks：How Realistic is The Threat [J]. In Proc. 2011 ACM WiSec，2011：47 – 52.

第 8 章　感知层入侵检测技术

网络安全问题主要来源于网络攻击，因此攻击检测是网络安全机制的重要任务。物联网感知层的设备通常部署在无人值守的环境中，并且是资源受限的，能源、计算能力、存储空间、通信距离的限制都使得传统的互联网入侵检测技术不能直接应用到物联网中，如何设计适用于物联网的入侵检测机制是物联网是否能够真正得到推广实用的关键问题之一。

8.1　物联网入侵检测技术概述

8.1.1　物联网入侵检测概述

传统的计算机网络中，检测代理通过观察网络流量、主机运行状态及系统日志来发现系统中的异常，从而检测入侵行为并采取相应措施。然而物联网和传统的计算机网络在终端设备类型、网络拓扑、数据传输等诸多方面都有极大的不同，面临的安全威胁也不尽相同，现有的技术成果已经不能解决物联网中的安全问题。因此，需要根据物联网自身的特点来设计相应的入侵检测方案。

物联网中，RFID 系统的入侵检测从目前现有的研究来看比较典型的检测模型有基于多决策树的入侵检测模型、基于有限自动机的入侵检测模型、基于免疫网络的入侵检测模型等，由于大部分入侵检测的工作都是在阅读器上进行的，而标签阅读器又往往与后台计算机系统密切相连，因此先后的 RFID 入侵检测机制很少考虑系统的资源和成本问题，与传统有线网络的入侵检测相似，因此本章不对 RFID 系统的入侵检测方法进行深入介绍。相反，物联网中的无线传感器网络由于往往部署在无人值守的环境中，与用户或后台通信受限，因此入侵检测系统对于网络安全而言就变得非常有必要，而且也出现了大批研究成果。本章重点对传感器网络入侵检测技术进行阐述。

8.1.2　常见的物联网入侵检测技术

国内外学者对物联网感知层的入侵检测技术进行了大量的研究，特别是针对无线传感器网络的研究，取得了一定的研究成果。从不同的角度对现有的检测技术进行分类，如图 8-1 所示。

与传统计算机网络一样，根据检测方法，可以分为误用检测和异常检测。误用检测技术观察系统内的行为，抽象出一组设定的签名，如 <IP 地址、权限、操作>，并根据专家知识维护一个包含恶意行为签名的规则知识库，将观察的行为签名与知识库中的规则进行匹配，如果匹配成功，则判定发生了入侵。这种检测方法可以高效地检测已知的攻击行为，但是在物联网环境中，新型攻击的出现需要频繁地更新知识库，并且随着时间的演进，知识库中的规则会越来越多，这对于存储空间有限的物联网感知设备是很大的挑战。异常检测认为任何一种偏离正常或期望模式的行为都是潜在的入侵，相对于误用检测，这种方法可以检测出未知的攻击，更具通用性，但是相应的误检率也会增加，常用的异常检测方法有统计分析

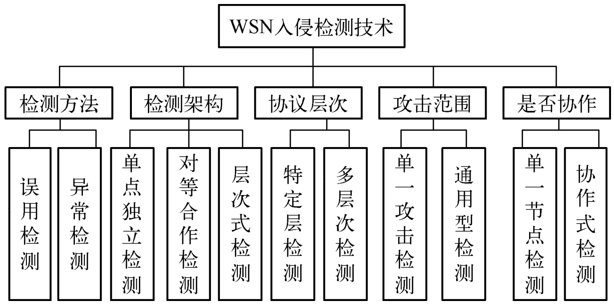


图 8-1　入侵检测技术分类

法、数据挖掘、人工智能、基于博弈论的检测、基于图的检测等。

根据检测架构，可以分为单点独立检测、对等合作检测和层次式检测。单点独立检测方式中，节点独立监测网络发现攻击，这种方式需要检测节点的通信及存储能力较强。对等合作检测方式中，网络中的每一个节点或部分节点上都部署检测程序，由节点自身处理一定的入侵行为，再通过合作的方式协作判定入侵。在物联网环境中，网络大都是异构的，即节点类型不完全相同，节点能力不同，担负的职责也不同，如在簇形分层网络中，簇头的能力高于成员节点，汇聚节点的能力高于簇头，层次式检测方式考虑了节点的差异性，从节省能耗的角度出发，在成员节点上部署简单的检测模块，成员节点以协作的方式为上层节点（簇头、汇聚节点）提供信息，检测及判决模块由上层节点来实现。

根据安全机制保护的协议层次，可以分为针对特定层的检测和针对很多层的检测。入侵检测系统（IDS）可以针对协议栈的某一层（现有的大部分研究针对网络层或应用层），也可以同时针对很多层。

根据攻击范围，可以分为针对单一攻击的检测和通用型检测。IDS 可以针对一组特定的攻击进行检测，也可以设计成通用的检测系统。

根据是否协作，可以分为单一节点检测和协作式检测。单一节点检测中，节点间没有信息的交互，仅处理自身获得的信息。在分布式和混合式架构中，节点可以通过合作检测入侵，或者直接向上层节点报告检测结果。

总体而言，大部分的攻击检测方案可以归为两类：集中式检测和邻居合作检测。

1. 集中式检测

集中式检测使用基站来检测攻击，如通过注入查询和收集相应信息来诊断传感器节点。大型的通信开销会导致潜在的检测失败，为了降低通信开销，他们的解决方案通过牺牲一些准确性来降低响应内爆。另一种集中式检测方法用来追踪故障节点。节点在它们的测量结果中增加了一些关于邻居的信息，并将这些信息传送到基站。这样基站就可以获取整个网络的拓扑。一旦基站获知了网络拓扑，则基于自适应路由更新消息，使用一个简单的分而治之的策略，就可以有效地跟踪故障节点。

2. 邻居合作检测

在邻居合作检测方法中，一个给定节点的邻居节点收集邻居的信息并且使用一个集体决策来检测攻击。Wang 等人在文章“On supporting distributed collab - oration in sensor networks”中提出一种分布式合作故障节点检测机制来使一个故障节点的邻居通过协作检测故

障。为了使邻居的通信有效，这种机制采用基于树的传播收集协议来从所有的邻居收集信息。这种收集方式具有低延时、低消息复杂度和低能耗的优点。Ding 等人在文章“Localized fault – tolerantevent boundary detection in sensor networks”中提出了另一种本地化的方法，该方法通过使用邻居的数据并使用统计学的方法对数据进行处理来检测故障传感器节点。阈值方法是一种特殊类型的邻居合作方法，如 Liu 等人在文章“On the intruder detection for sinkhole attack in wireless sensor networks”中引入一种新的邻居合作方法来检测内部攻击。该算法的特色在于不需要正常节点或恶意节点的先验知识，这对动态攻击行为是很重要的。此外，他们的算法可以用来检查网络活动的任何方面，同时评估多个属性。

3. 其他检测方案

尽管许多攻击检测机制使用集中式检测方法或邻居合作/本地化方法来监测节点的活动，但是还存在一些其他检测方案：代码测试方案和位置验证方案。

1）代码测试方案

代码测试方案包括两种：基于软件的方案和基于硬件的方案。

（1）基于软件的方案

基于软件的方案依靠最佳的程序代码和精准的时间管理。通过引入一个最优方案的验证过程来实现基于软件的认证，该验证过程验证传感器节点的存储内容（存储内容是通过计算随机选择的存储器区域的哈希值获得的）。

（2）基于硬件的方案

一般基于硬件的方案是基于公钥加密的，并且需要额外的计算能耗，像传输大量信息一样，使得这些方案不适用于能量有限的无线传感器网络。Krauss 等人在文章“Detecting node compromise inhybrid wireless sensor networks using attestation techniques”中提出了一些簇节点比大多数的簇拥有更多的资源，并在混合无线传感器网络中配备了可信平台模块。其基于硬件的认证协议使用配备了可信平台模块的节点作为信任锚节点，并能更有效地使能认证。然而，他们的机制只是在混合无线传感器网络中有意义。

2）位置验证方案

Song 等人在文章“Sensor node compromisedetection：the location perspective”中提供了一种通过对比节点先前位置和当前位置的检测节点妥协的方案。该方案的主要思想是基于假定节点的妥协通常由三个阶段组成：物理获取并影响传感器，重新部署妥协的传感器，妥协的传感器节点重新入网后发动攻击。在一些应用中攻击者可能不能准确地将妥协传感器部署到其原来的位置，他们的方案在妥协节点更换位置或身份时可以检测到妥协事件。但是有时恶意攻击者可以通过和节点通信、违反节点的安全机制、不物理接触节点或移动节点位置的方式来使节点妥协。在这些情况下，这种机制无法检测出妥协事件。

8.2　通用型入侵检测算法

8.2.1　基于分簇的入侵检测算法

C. Loo 等人在文章“Intrusion detection for routing attacks in sensor networks”中使用分簇

算法实现对路由攻击的检测，每个节点检测其接收到的路由信息。在一个检测时间内，每个传感节点根据具体的特征将路由记录进行分类，每一个特征向量可以看成多维特征空间中的一个点。文章假设攻击流量远小于正常的网络流量，并且在统计信息上与正常的流量不相同，在特征空间上呈现异常。

异常检测算法包括以下两个阶段。

① 训练阶段：节点每隔一段时间采集一次数据，共采集 N_{Tr} 次，并生成样本集合 $C_{Tr}=\{c_m, m=1,2,\cdots,N_{Tr}\}$。每个数据样本由12个特征属性组成，主要描述节点的路由和流量状况。使用固定宽度分簇算法对样本采集 C_{Tr} 进行分簇，得到 C_{Tr} 的子集合，分簇集合 $\Phi=\{\phi_1,\phi_2,\cdots,\phi_s\}(1\leqslant s\leqslant N_{Tr})$。然后对分簇集合中的每个簇 $\phi_v(1\leqslant v\leqslant s)$，计算其内部包含的样本个数 $|\phi_v|$ 与总样本点的比值 N_{Tr}，若 $|\phi_v|/N_{Tr}$ 小于给定的阈值 τ，则 ϕ_v 标记为异常类。

② 测试阶段：将待测试的网络流量样本 C_n 置于簇集合中进行比较，如果 C_n 与异常类 ϕ_v 的距离 $|c_n-\phi_v|$ 小于报警阈值 ω，则说明测试点 C_n 属于此异常类 ϕ_v，入侵检测系统将标记 C_n 为异常点，并产生报警信息。采用簇的异常流量监测算法不依赖于有标记的训练集合，能够检测出未知的攻击，可以应用在大多数路由攻击检测中，特别是对流量影响比较大的检测攻击检测中。

S. Rajasegarar 等人研究了在分层网络中使用分簇算法实现入侵检测。在中心化检测方法中，传感节点收集数据发送给网关，由网关节点进行检测，这种方法造成了大量的通信开销。该方案中，每个节点同样使用固定宽度分簇算法进行本地分簇，然后节点发送统计信息给自己的父节点。这种通信不需要大量的开销，父节点收到分簇信息后，对信息进行聚合和去除重复的簇，将处理后的分簇信息发送给网关节点，网关节点再根据分簇信息进行异常检测。具体检测过程如下。

首先对输入的数据进行归一化，对于给定的数据集合 $v_{kj}, k=1,\cdots,m$，将其转换成 $u_{kj}=(v_{kj}-u_{vj})/\delta_{vj}$，其中向量 u_{vj} 和 δ_{vj} 为第 j 个属性的均值和方差。然后将 u_{kj} 归一化到［0，1］区间内：$\overline{u_{kj}}=(u_{kj}-\min u_j)/(\max u_j-\min u_j)$。

一个普通的成员节点 s_i 收集数据 x_i，s_i 将本地的正常模式发送给其父节点：

$$\left(\sum_{k=1}^{m} x_k^i,\ \sum_{k=1}^{m}(x_k^i)^2, m, x_{\max}^i, x_{\min}^i\right)$$

其中 m 表示 $|x_i|$。

父节点计算全局正常模式 $(\mu_G,\delta_G^2,\chi_{\max}^G,\chi_{\min}^G)$，并将全局的正常模式返回给成员节点。收到全局正常模式之后，每个成员节点再利用固定宽度分簇算法进行本地检测，如果样本点与簇中心的欧几里得距离大于一个用户固定的阈值 ω，则产生一个以该样本点为中心的新簇。为了降低簇的数量，使用内部簇距离对不同的簇进行聚合。对于簇 c_1 和 c_2，如果它们的距离 $d(c_1,c_2)$ 小于 ω 就将进行聚合。最后使用 k 邻居（KNN）分簇算法进行异常检测。令 ICD_i 为簇 i 的内部簇距离（KNN），AVG(ICD) 和 SD(ICD) 为所有的内部簇距离的均值和方差，如果：$\mathrm{ICD}_i>\mathrm{AVG(ICD)}+\mathrm{SD(ICD)}$，则簇为异常簇。

Rajasegarar 等人在文章“Distributed anomaly detection in wireless sensor networks”中采用了基于四分之一球的 SVM 实现分层传感器网络的异常检测。网络分层模型如图 8-2（a）所示。

首先，每个普通成员节点计算本地的 radii，然后簇头收集本地计算的 radii，计算全局的

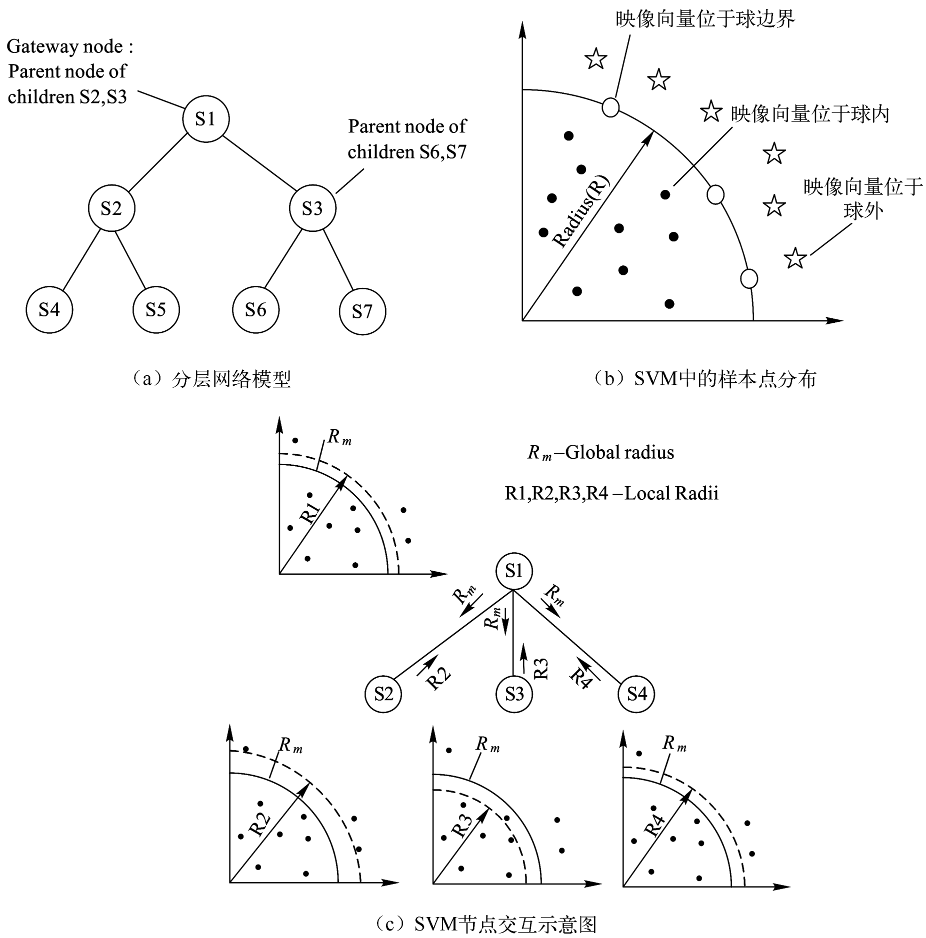
半径，依次建立正常模式。然后普通成员节点根据正常模式检测异常，使用优化问题求解问题解决正常模式的建立。

优化问题形式化表示为：

$$\min_{R\in\Re,\varepsilon\in\Re^n} R^2+\frac{1}{vn}\sum_{i=1}^{n}\xi_i,\quad \|\varphi(x_i)\|^2\leqslant R^2+\xi_i,\ \xi_i\geqslant 0$$

其中 x_i 是数据向量，对应的向量 $\varphi(x_i)$ 称为映象向量，R 为四分之一球的半径，$\{\xi_i:i=1,\cdots,n\}$ 是松弛变量，允许一定的映象向量落在四分之一球的外面。这个问题可以通过拉格朗日算法求解。如图 8-2（b）所示，映象向量最终可以分为三类：位于球内、位于球的边界上或位于球外。然后，簇头节点收集节点本地计算的 radii 以获得全局 R_m。可以通过多种方法计算 R_m，如均值、中间值、最小值、最大值。当普通成员节点收到 R_m 之后，初始化检测引擎。如果测试样本 x_i 满足 norm $\tilde{k}(x_i,x_i)>R_m^2$，就被确认为异常。

这种方法的缺点是需要大量的数据处理过程，这也是由 SVM 的自身复杂度决定的。但是在检测过程中，只有一个半径参数需要节点间交互，见图 8-2（c），节省了通信开销。



（a）分层网络模型

（b）SVM中的样本点分布

（c）SVM节点交互示意图

图 8-2 基于 SVM 的入侵检测

8.2.2 基于博弈论的入侵检测算法

Agah 等人在文章 “Intrusion detection in sensor networks: a non – coorperative game approach” 中将博弈论中的非合作模型应用到传感器网络的入侵检测中，并提出了一种新的解决方案，即利用两人非零合作和动态博弈理论对入侵者和防护系统之间的对抗行为进行分析，建立入侵监测和响应模型，从而制定出一个防御策略，以帮助提高入侵检测成功率。

算法将博弈论监测机制应用在分簇的传感器网络中，并按照博弈论建立传感器网络攻防双方的模型，对于传感器网络中给定的簇头节点，使用二元组 {AS，SS} 分别表示攻击者和检测系统的策略空间。其中，攻击者有三个策略，即 AS_1 攻击簇、AS_2 不对任何簇进行攻击和 AS_3 攻击其他某个簇；检测系统有两个策略，即 SS_1 保护簇和 SS_2 保护某个簇。可用矩阵 A 表示检测系统的收益值和矩阵 B 表示攻击的收益值：

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{12} \\ a_{21} & a_{22} & a_{23} \end{bmatrix},\ B=\begin{bmatrix} b_{11} & b_{12} & b_{12} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$$

矩阵中的各元素表示攻防双方采用不同策略对时所对应的收益值。例如，a_{11} 表示用（AS_1，SS_1）策略对进行攻击和保护时检测系统的收益值。

可以根据此博弈模型分析攻防博弈的动态过程：网络攻击者根据自己的经验和对网络状态信息的判断，选择攻击或放弃；网络探测器监测网络数据，向检测系统报告自己的检测结果，检测系统权衡其可能的收益并做出决策。可以证明，当网络攻防双方采用策略对（AS_1，SS_1）时，该博弈模型达到纳什均衡。因此，对于检测系统来说，最好的策略就是找到最合适的簇头节点进行防范。基于博弈论的检测算法有利于深入分析入侵检测结果，权衡检测效率和网络资源，得出最终合理的响应决策。但是检测时需要人工干涉，系统的自适应能力比较差。

8.2.3 基于模糊理论的阻塞攻击入侵检测算法

Misra 等将模糊理论应用到信息战中的阻塞攻击检测中，并在文章 “Information Warfare – Worthy Jamming Attack Detection Mechanism for Wireless Sensor Networks Using a Fuzzy Inference System” 中提出了一种基于模糊推理的检测方案。

该文章首先分析了下面 9 个检测指标。

载波侦听时间（Carrier Sensing Time，CST）：为了避免数据包碰撞，MAC 层协议（如 BMAC）提供了冲突避退机制，在节点开始传输数据之前，需要首先侦听信道并持续一个预定义的时间，如果信道忙碌，节点需要等待一个特定的时间再重新尝试，从节点准备发送数据到数据包发送出去或放弃的时间间隔称为载波侦听时间。这个指标需要 MAC 层提供冲突避退机制，不能检测物理层能量攻击，同时需要确定一个阈值来确定是否异常。

包发送率（Packet Send Ratio，PSR）：一段时间内节点发送的数据包数目除以节点正常需要发送的数据包数目即为 PSR。计算 PSR 时需要一个预置的时间，用来计算节点需要发送的数据包数目。

包传送率（Packet Delivery Ratio，PDR）：一般由发送方衡量，即包成功传输的概率，

因为需要类似于 TCP 的可靠传输，作者认为这在资源受限的传感网络是浪费资源的。

坏包率（Bad Packet Ratio，BPR）：类似于 PDR，由接收方衡量，指的是 CRC 校验不通过的数据包。

接收信号强度标准差（Standard Deviation in Received Signal Strength，SDRSS）：首先采集正常信号的强度，计算一段时间内的额信号标准差 σ。然后监测异常信号强度与正常信号之间的平均偏差 $\bar{d}$，如果 $\bar{d}\leqslant\sigma$，就没有阻塞信号，否则发生阻塞。作者认为这种方法并不适合 WSN，因为：

① 如果阻塞攻击者节点与正常信号强度相同，则无法检测；

② 需要节点级的采样；

③ 节点计算负担大。

误码率（Bit Error Rate，BER）：接收方计算一个完整的比特流中有多少比特的错误，可以用来检测响应式阻塞，但是这种方法对节点的负担很大。

接收信号强度（Received Signal Strength，RSS）：一般不单独作为一个衡量指标，但是可以与其他指标（如噪声强度或 BER）混合来检测阻塞。

信噪比（Signal－to－Noise Ratio，SNR 或 Signal－to Jammer Power Ratio，SJR）：接收信号的强度与噪声强度的比值，从物理层上检测阻塞。其他从链路层检测攻击的指标如 PDR、BPR、BER 等应该配合 SNR 来全面地验证检测结果。

能量消耗量（Energy Consumption Amount，ECA）：因为阻塞使得节点长期处于避退期，而不能进入 IDLE 模式，从而造成了能量耗尽。这种方法存在两个缺点：

① 在正常高负载的情况下，能耗的阈值也是偏高的；

② 对于阻塞者不通过占用信道攻击时，能量的变化并不明显。

基于上述分析，最终选择了 SNR 和 BPR 作为检测指标。该文章将 BPR 称为 PDPT（Packets Dropped per Terminal），即计算在一个仿真周期中的 BPR。

如果 X 为一组对象的集合，称为由 x 产生的论域（Universe Of Discourse，UOD），X 中的模糊集合 A 定义为：$A=\{(X,\mu_A(x)):x\in X\}$。其中，$\mu_A(x)$ 称为模糊集合 A 的隶属度函数。隶属度函数将元素 X 映射到区间为［0，1］的隶属值上。

模糊逻辑可分为三个步骤：模糊化、模糊推理及去模糊化。

1. 模糊化

模糊推理系统包括两个输入 SNR 和 PDPT，并采用梯形模糊函数作为隶属度函数。隶属度函数如下：

$$\mu_{\text{set}}(\text{uod})=\begin{cases}\dfrac{\text{uod}-a}{b-a}, & a\leqslant\text{uod}\leqslant b\\ 1, & b<\text{uod}<c\\ \dfrac{d-\text{uod}}{d-c}, & c\leqslant\text{uod}\leqslant d\\ 0, & \text{otherwise}\end{cases}$$

系统输出值为 JI（Jamming Index）。JI 的变化范围为 0～100，分别代表“没有阻塞”和“完全阻塞”。具体参数见表 8-1，参数通过两个步骤获得：

① 由专家经验确定；

② 通过检测结果的回馈进行修正。

隶属度函数的图形表示如图 8-3 ~ 图 8-5 所示。

表 8-1　隶属度函数中的变量值

论域（uod）	集合	*a*	*b*	*c*	*d*
SNR	LOW	-0.5	0	1	1.5
	MEDIUM	1	1.5	10	12
	HIGH	10	12	3900	4000
PDPT	LOW	-5	0	10	15
	MEDIUM	10	15	25	30
	HIGH	25	30	50	55
JI	LOW	-5	0	25	30
	MEDIUM	50	55	75	80
	HIGH	75	80	100	105

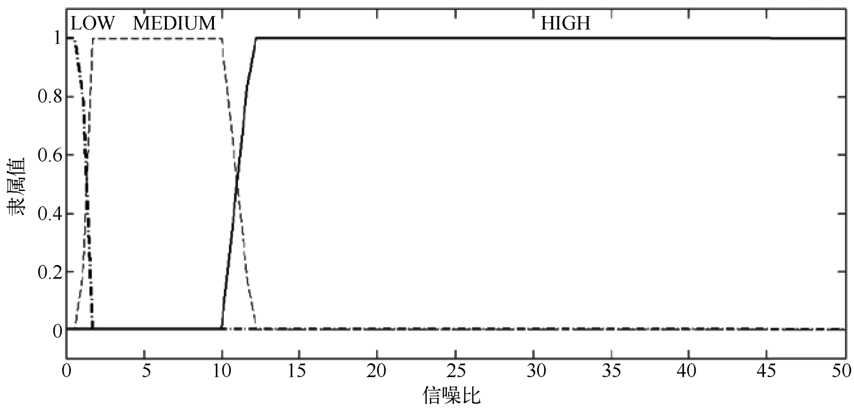


图 8-3　SNR 隶属度函数

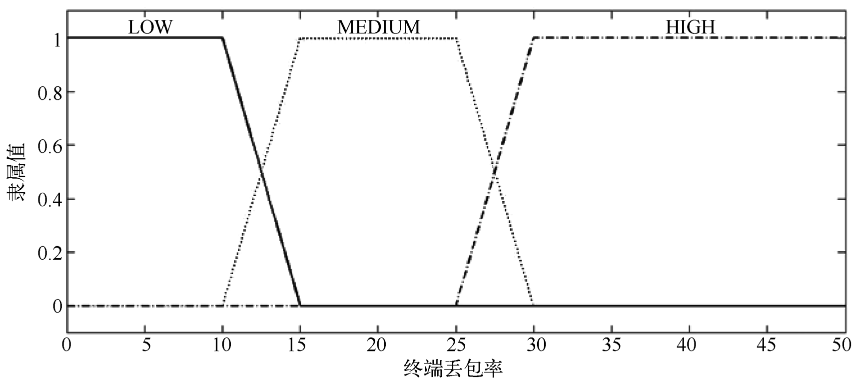


图 8-4　PDPT 隶属度函数

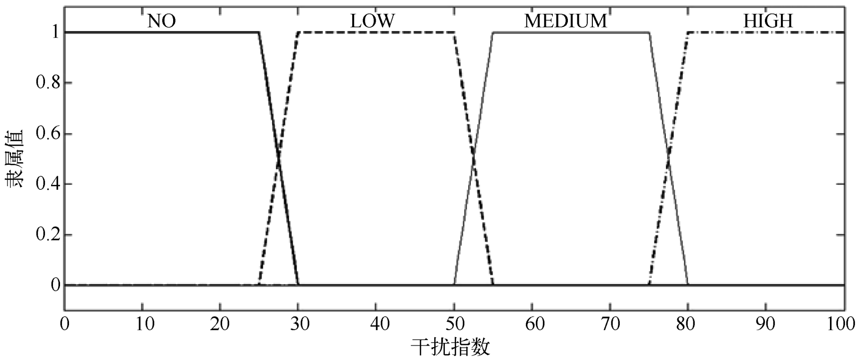


图 8-5　JI 隶属度函数

2. 模糊推理

模糊逻辑的第二个步骤是模糊推理，在这一步骤中，通过专家经验制定一系列的规则，并通过系统的输出进一步优化。这些规则包括：

If SNR is LOW and PDPT is LOW then JI is HIGH.
If SNR is LOW and PDPT is MEDIUM then JI is HIGH.
If SNR is LOW and PDPT is HIGH then JI is HIGH.
If SNR is MEDIUM and PDPT is LOW then JI is LOW.
If SNR is MEDIUM and PDPT is MEDIUM then JI is MEDIUM.
If SNR is MEDIUM and PDPT is HIGH then JI is HIGH.
If SNR is HIGH and PDPT is LOW then JI is NO.
If SNR is HIGH and PDPT is MEDIUM then JI is LOW.
If SNR is HIGH and PDPT is HIGH then JI is MEDIUM.

3. 去模糊化

$$\mathrm{JI} = \mathrm{COA} = \frac{\sum_{\mathrm{uod}=a}^{b} \mu_{\mathrm{set}}(\mathrm{uod}).\ \mathrm{uod}}{\sum_{\mathrm{uod}=a}^{b} \mu_{\mathrm{set}}(\mathrm{uod})}$$

如上式所示，去模糊化采用重心法（Centroid of Area，COA），根据计算的结果，基于模糊逻辑，通过 2 – 均值分簇算法来确认检测攻击，检测过程如下。

① 根据信息战场景，确定 JI 的下限值 LC，JI≥LC 的节点为被阻塞的节点。

② 将 t 个被阻塞的节点号放入一张列表中。

③ 对每一个被阻塞的节点：

步骤 1，识别并计算一跳邻居的数目 n；

步骤 2，遍历其 n 个邻居，计算在阻塞节点列表中邻居节点的数目 nj，对应的邻居节点集合称为阻塞邻居簇（jammed neighbors cluster）。

步骤 3，没有在阻塞节点列表中的邻居数目为 $n-nj$，未阻塞邻居节点集合为非阻塞邻居簇（non – jammed neighbors cluster）。

步骤 4，如果 $nj > n/2$，则大部分邻居节点被阻塞，从而确认该节点也被阻塞；

步骤 5，如果 $nj \leqslant n/2$，则需要进一步的确认。

● 计算阻塞邻居簇的阻塞指数均值：

$$\overline{jij} = \frac{\sum_{k=1}^{nj} ji_k}{nj}$$

● 计算非阻塞邻居簇的阻塞指数均值：

$$\overline{jinj} = \frac{\sum_{k=1}^{n-nj} ji_k}{n - nj}$$

● 计算阻塞邻居簇的中心坐标：

$$\left(\overline{x_j} = \frac{\sum_{k=1}^{nj} ji_k \cdot x_k}{\sum_{k=1}^{nj} ji_k}, \overline{y_j} = \frac{\sum_{k=1}^{nj} ji_k \cdot y_k}{\sum_{k=1}^{nj} ji_k}\right)$$

● 计算非阻塞邻居簇的中心坐标：

$$\left(\overline{x_{nj}} = \frac{\sum_{k=1}^{n-nj} ji_k \cdot x_k}{\sum_{k=1}^{n-nj} ji_k}, \overline{y_{nj}} = \frac{\sum_{k=1}^{n-nj} ji_k \cdot y_k}{\sum_{k=1}^{n-nj} ji_k}\right)$$

● 计算衡量节点和阻塞邻居簇的中心的距离的平方：

$$d_j^2 = (x - \overline{x_j})^2 + (y - \overline{y_j})^2$$

● 计算衡量节点和非阻塞邻居簇的中心的距离的平方：

$$d_{nj}^2 = (x - \overline{x_{nj}})^2 + (y - \overline{y_{nj}})^2$$

如果：

$$(\overline{jij}/d_j^2) \geqslant (\overline{jinj}/d_{nj}^2)$$

则确认节点被阻塞，否则节点未被阻塞，同时将节点从阻塞节点列表中删除。

8.2.4　基于人工免疫的入侵检测技术

由于传感器网络缺乏人工干预及资源受限的特点，使得无线传感器网络的入侵检测面临着巨大的挑战。近年来生物免疫启发的人工免疫系统（AIS）由于其在工作原理上与入侵检测的一致性得到了广泛的关注，AIS 的基本特征包括自组织、分布式、高鲁棒性、轻量级、多层次及多样性等。这些特征使其在入侵检测方面体现出了优势，并取得了很多成果。受生物领域中危险理论的启发，“基于危险理论的无线传感器网络入侵检测模型”一文提出了一种基于危险理论的入侵检测模型，并通过仿真实验验证了提出的模型在检测率、误检率和能量消耗方面具有优势。

危险理论是自然免疫学科中的最新研究成果，危险模型如图 8-6 所示。危险模型在细胞和信号的基础上增加了额外的一层，认为 APC 由受难细胞（如受到病原体侵入、毒素侵入、创伤等影响的细胞）发出的危险信号触发，危险信号被 APC 识别，是引起免疫应答的关键因素。图 8-6 描述了危险模型中免疫应答的响应示意图。免疫过程可分为以下步骤：

一个非正常死亡的细胞发出了一个危险信号，邻近的抗原提呈细胞 APC 被激活并开始识别和捕获抗原，APC 通知本地的淋巴结并把所识别的抗原提呈给淋巴细胞，淋巴细胞产生抗体进行抗原识别。

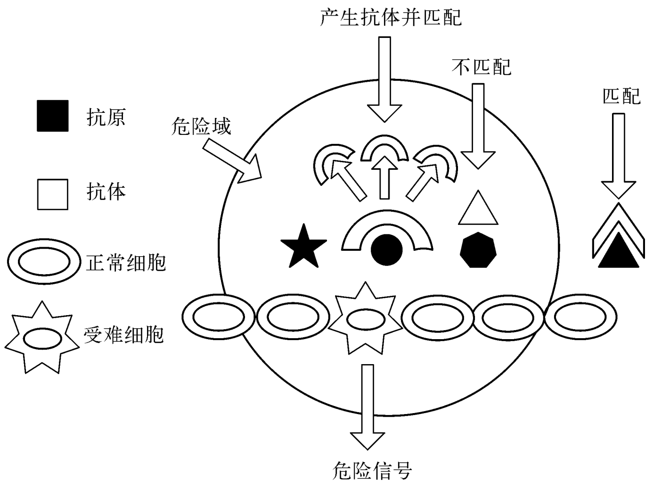


图 8-6　危险模型示意图

从本质上讲，危险信号的产生将会在受难细胞周围建立一个危险区域，在危险区域内的淋巴细胞才会被激活，产生大量的对应能匹配该抗原的抗体。而那些不在危险域内的淋巴细胞则不被激活，因而也不能产生抗体。

危险模式应用于无线传感网络的入侵检测主要存在以下两个优点：

① 发生危险时才会触发检测过程，可以降低误检率并且降低不必要的能量消耗；

② 危险域可以根据不同的危险程度或安全策略来确定，可以提高检测系统的灵活性。

根据危险理论，受难细胞发出危险信号，并在其周围建立一个危险域，其中的抗原被 APC 捕获，APC 提呈抗原提供共同刺激信号，从而引起免疫应答。

利用危险理论的工作原理，将检测过程分成三个阶段：危险感知阶段、抗原提呈阶段和决策阶段，图 8-7 描述了之前提出的基于危险理论的无线传感器网络入侵检测模型。

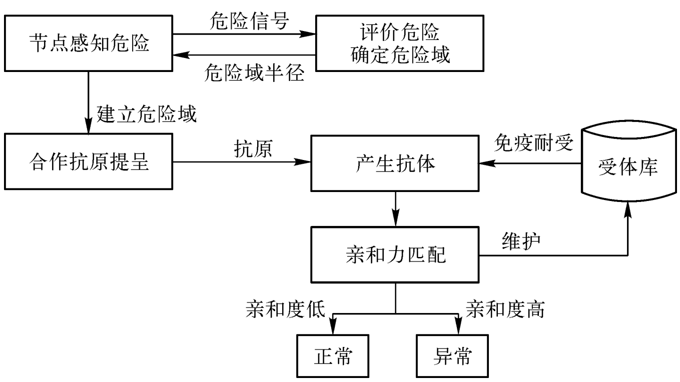


图 8-7　入侵检测模型

从能源的角度来看，每个节点同时运行一个完整的检测实例是不合适的。因此，该模型利用 IEEE 802. 15. 4 的分层结构实现合作式检测，边界终端节点只检测自身属性的变化来感

知风险，中心节点（称为决策节点）接收到危险信号后确定危险程度及危险域的大小并要求获得抗原，危险域内的节点合作提供网络流量信息形成抗原，中心节点负责产生抗体，动态地维护抗体，并将抗体与抗原进行匹配确定是否发生了入侵。

表 8-2 给出了基于危险理论的 AIS 到上述系统模型的隐喻。

表 8-2　危险理论 AIS 到系统模型的隐喻

危险理论 AIS	系统模型
淋巴细胞	检测节点和决策节点
病原体	入侵节点或具有传播性的蠕虫、病毒
凋亡细胞	由于能量耗尽等原因不工作的节点
受难细胞	由于受到入侵影响，感知到危险的节点
危险信号	发送给决策节点表明存在潜在入侵的信息
APC	危险域中的节点合作提供抗原
抗原（Antibody）	危险域中的全局信息
抗体（Antibody）	由耐受过程产生的异常模式
受体（Receptor）	抗原和抗体都由受体构成，每个受体描述每个节点流量信息的正常模式

1. 危险感知

在生物免疫系统中，如果一个细胞由于正常原因死亡（凋亡细胞），细胞实体在分解之前就会被清除。由于非正常原因坏死的细胞（受难细胞）就会分解实体并且释放出危险信号。类似地，在该模型中，节点感知自身的变化发现危险并释放出危险信号，通常情况下，节点能够在失去工作能力之前发现危险并且做出反应。

节点自身属性的异常变化反映了潜在的危险。节点的物理层和媒体接入层的属性都可以在本地获得，因此应关注这两层的属性信息。用来感知危险的属性用 DFi（Danger Features）来表示。

DF1：电源能量下降速率。传感节点是资源受限的，对 DoS 类攻击非常敏感。DF1 的计算式为

$$\mathrm{DF1} = C_{\mathrm{power}}/\Delta t$$

其中，C_{power}表示在 Δt 时间内电能的变化量。

DF2：数据包发送回退频率。由于采用了冲突避退机制，当发生包阻塞攻击时，这个属性的变化明显。DF2 的计算式为

$$\mathrm{DF2} = \sum_{t}^{t+\Delta t} N_{\mathrm{BP}}/\Delta t$$

其中，$\sum_{t}^{t+\Delta t} N_{\mathrm{BP}}$ 表示在 Δt 时间内回退的数据包个数。

DF3：平均回退持续时间。此属性的变化可以发现持续的阻塞干扰类攻击。DF3 的计算式为

$$\mathrm{DF3} = T_{\mathrm{BP}}/\Delta t$$

其中，T_{BP}表示在 Δt 时间内总共回退等待的时间。

DF4：ACK 成功率。在发送数据之后，节点通常希望获得 ACK 以证实数据发送成功，

ACK 成功率过低也表明存在危险。DF4 的计算式为

$$\mathrm{DF4} = \sum_{t}^{t+\Delta t}(N_{\mathrm{ACK}}/N_{\mathrm{SP}})$$

上式计算了在 Δt 时间内发送的数据包的个数与实际收到的 ACK 的个数的比值。

DF5：数据帧接收频率。接收到的数据帧频率的异常变化暗示着危险，如节点作为攻击目标时，接收到的数据帧数目增大，接收频率增大。DF5 的计算式为

$$\mathrm{DF5} = \sum_{t}^{t+\Delta t} N_{\mathrm{RF}}/\Delta t$$

DF6：数据帧发送频率。发送的数据帧数频率的异常变化也暗示着危险，如发生大规模的蠕虫或阻塞攻击时，节点通常要转发这些恶意的数据包导致发送数据帧数目增大，发送频率增大。而发生 Sinkhole 攻击时，本来作为正常路由的节点将不再转发数据导致发送数据帧数目骤减，发送频率骤减。DF6 的计算式为

$$\mathrm{DF6} = \sum_{t}^{t+\Delta t} N_{\mathrm{SF}}/\Delta t$$

将每个属性归一化，并给定统一的变化阈值 δ，在 t 时刻，如果 $\mathrm{CF}i = |\mathrm{DF}i_t - \mathrm{DF}i_{t-1}| > \delta$，则认为属性 DF$i$ 发生了不正常的变化，可能存在危险。

感知到危险之后，节点发送危险信号给决策节点，危险信号表示为

$$\mathrm{DS} = <\mathrm{Timestamp}, \{(\mathrm{DF}i, \mathrm{CF}i)\}>$$

危险感知过程可以利用每次节点的正常工作时间，不需要产生额外的调度将节点唤醒。

2. 抗原提呈

一旦决策节点接收到危险信号，便要建立一个危险域，危险域以发出危险信号的节点为中心，覆盖范围称为危险域半径，以跳数为单位。危险域半径与危险程度有关，危险程度表示为

$$D_{\mathrm{danger}} = \sum_{j=0}^{nd} \sum \omega i_j * \mathrm{CF}i_j$$

其中，nd 为一个时间段内决策节点接收到的危险信号的个数；ωi 为每一个危险属性变化的权重。危险域半径为

$$Ra = \lceil s \times D_{\mathrm{danger}} \rceil$$

参数 s 为保护的无线网络的安全等级，从 Ra 的表达式可以看出，危险半径与网络的危险程度和网络的安全等级成正比。

感知到危险的节点在自己的 Ra 跳范围之内建立危险域，此节点向危险域内的节点广播流量日志获取请求。在一些情况下，如蠕虫攻击，很多节点感知到危险，危险域就会存在重叠，这种情况下，节点选择最近的节点上传自己的流量日志。流量日志表示为 $\log = <Ps, Pr, Pf>$，其中：

$$\begin{cases} Ps = N_{\mathrm{SP}}/\Delta t \\ Pr = N_{\mathrm{RP}}/\Delta t \\ Pf = N_{\mathrm{FP}}/\Delta t \end{cases}$$

N_{SP}、N_{RP} 和 N_{FP} 分别表示节点在 Δt 时间内发送、接收和转发的网络数据包的数目。在决策节点接收到危险域内所有节点的流量日志或等待超时之后，决策节点停止收集，并提呈抗原。

受体是组成抗原和抗体的基本单元，每一个节点 i 都有一个 Id_i 和相对应的受体，受体

表示为

$$R(\mathrm{Id}_i) = <\mathrm{Id}_i, Ps, Pr, Pf, \{\mathrm{NeighborList}\}>$$

假设危险域有 k 个节点，抗原可以表示为

$$Ag = \{\bigcup_{i=1\sim k} R(\mathrm{Id}_i)\}$$

3. 决策

决策节点负责分析提呈的抗原，确认入侵行为的存在。

分析过程采用传统的自我－非我识别，通过计算抗原和抗体之间的亲和度确认是否发生了入侵。抗原通过对受体库进行免疫耐受过程产生。受体库为每个节点预定义的自我集合，仅存储在决策节点上。APC 激活了受体库，为每个节点提供数目为 m 的非我受体，非我受体组成抗原。

为了区别用于组成抗原和抗体的受体，使用$\hat{R}(\mathrm{Id_i})$来表示耐受产生的非我受体。抗体表示为

$$Ab = \{\bigcup_{i=1\sim k} \hat{R}(\mathrm{Id}_i)\}$$

从上式可以看出，受体是组成抗原和抗体的基本单元，也是用来识别的基本单元。决策节点从抗原中提取出 Id_i，并激活受体库产生抗原。如图 8-8 所示，分别将抗原与抗体中的受体使用亲和度函数表示。有文献使用 Euclidean 距离函数来计算亲和力，对每个 Id_i，受体之间的距离为

$$A(\mathrm{Id}_i) = \sqrt{(ps-\hat{ps})^2 + (pf-\hat{pf})^2 + (pr-\hat{pr})^2}$$

抗原和抗体之间的亲和力为

$$A(Ag, Ab) = \frac{1}{\sum_{i=1}^{k} A(\mathrm{Id}_i)}$$

如果 $A(Ag,Ab) > \beta$，则认为确实发生了入侵。β 为亲和力阈值。

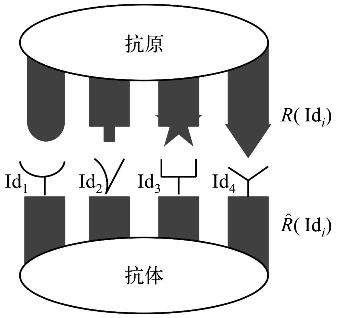


图 8-8　抗原、抗体亲和力匹配

图 8-8 描述了危险域中有 4 个节点的抗原和抗体的匹配过程，从图中可以看出 $R(\mathrm{Id}_1)$、$R(\mathrm{Id}_4)$与$\hat{R}(\mathrm{Id}_1)$、$\hat{R}(\mathrm{Id}_4)$匹配度高，这就说明 Id_1 及 Id_4 节点为潜在的入侵者和严重受害者。

自我受体库不应是静态的，应是随着检测结果动态变化的，如可以通过记忆受体减低检测试验或淘汰长时间没有用到的受体等，该模型关注整个检测模型的性能，对自我受体库的维护不做过多的讨论。

参考文献

[1] Kocher P, Jaffe J, Jun B. Differential power analysis. Lecture Notes in Computer Science, 1999, 1666, 388－397.

[2] Messerges T. S., DabbishE. A., Sloan R. H. Power analysis attacks of modular exponentiation in smartcards. In Proceedings of the First International Workshop on Cryptographic Hardware and Embedded Systems CHES'

99, Worcester, MA, USA, August 1999: 144－157.

[3] Thanassis Giannetsos, Tassos Dimitriou, Neeli R. Prasad. Self－Propagating Worms in Wireless Sensor Networks. Co－NEXTStudent Workshop'09 Proceedings of the 5th international student workshop on Emerging network experiments and technologies: pp. 31－32. ACM New York, NY, USA.

[4] QijunGu, Rizwan Noorani. Towards Self－propagate Mal－packets in Sensor Networks. Proceedings of the first ACM conference on Wireless network security WiSec'08: 172－182.

[5] Ana Paula R. da Silva, Marcelo H. T. Martins, Bruno P. S. Rocha, Antonio A. F. Loureiro, Linnyer B. Ruiz, Hao Chi Wong. Decentralized Intrusion Detection in Wireless Sensor Networks. Proceedings of the 1st ACM international workshop on Quality of service & security in wireless and mobile networks,, NY, USA 2005: 6－23.

[6] Yi Ping, Wu Yue, Chen Jialin, Towards an artificial immune system for detecting anomalies in wireless mesh networks. China Communication. 2011. 5, 3: 116－126.

[7] Shun－Sheng Wang, Kuo－Qin Yan, Shu－Ching Wang, Chia－Wei Liu. An Integrated Intrusion Detection System for Cluster－based Wireless Sensor Networks. Expert Systems with Applications, 2011, 38: 15234－15243.

[8] E. Ngai, J. Liu, and M. Lyu, "On the Intruder Detection for Sinkhole Attack in Wireless Sensor Networks," ICC'06, Istanbul, Turkey, June 2006.

[9] C. Loo et al., "Intrusion Detection for Routing Attacks in Sensor Networks," Int'l. J. Distrib. Sensor Networks, vol. 2, no. 4, 2006: 313－32.

[10] S. Rajasegarar et al., "Distributed Anomaly Detection in Wireless Sensor Networks," Proc. 10th IEEE Int'l. Conf. Commun. Systems, Singapore, Oct. 2006.

[11] Rajasegarar S, et al. Quarter sphere based distributed anomaly detection in wireless sensor networks. Presented at the IEEE international conference on communications, June 2007.

[12] Agah A, Das S K, Basu K. "Intrusion detection in sensor networks: a non－coorperative game approach". Proc of the 3rd IEEE International Aymposium on Network Computing and Application, Boston, 2004: 343－3456.

[13] Sudip Misra, Ranjit Singh and S. V. Rohith Mohan "" Information Warfare－Worthy Jamming Attack Detection Mechanism for Wireless Sensor Networks Using a Fuzzy Inference System. Sensors. 2010, 10: 3444－3379.

[14] 傅蓉蓉，郑康峰，芦天亮，杨义先．基于危险理论的无线传感器网络入侵检测模型，通信学报［J］，2012. 9: 31－37.

[15] C. Jaikaeo, C. Srisathapornphat, and C.－C. Shen, "Diagnosis of sensornetworks," in Proc. IEEE International Conf. Commun., 2001, (5): 1627－1632.

[16] J. Staddon, D. Balfanz, and G. Durfee, "Efficient tracing of failednodes in sensor networks," in Proc. 1st ACM International Workshop Wireless Sensor Networks Applications, 2002: 122－130.

[17] G. Wang, W. Zhang, and G. Cao, "On supporting distributed collab－oration in sensor networks," in Proc. IEEE Military Communi. Conf., 2003, (2): 752－757.

[18] S. Marti, T. Giuli, K. Lai, and M. Baker, "Mitigating routing misbe－havior in mobile ad hoc networks," in Proc. 6th Annual International Conference Mobile Computing Networking, 2000: 255－265.

[19] M. Ding, D. Chen, K. Xing, and X. Cheng, "Localized fault－tolerantevent boundary detection in sensor networks," in Proc. IEEE INFO－COM, 2005, (2): 902－913.

[20] B. Krishnamachari and S. Iyengar, "Distributed Bayesian algorithmsfor fault－tolerant event region detection in wireless sensor networks," IEEE Trans. Comput., 2004, (53): 241－850.

[21] T. Palpanas, D. Papadopoulos, V. Kalogeraki, and D. Gunopulos, "Distributed deviation detection in sensor networks," ACM SIGMODRecord, 2003, (32): 77 – 82.

[22] F. Liu, X. Cheng, and D. Chen, "Insider attacker detection in wirelesssensor networks," in Proc. IEEE INFOCOM, 2007: 1937 – 1945.

[23] A. Seshadri, A. Perrig, L. v. Door n, and P. Khosla, "SWATT: Soft Ware – based ATTestation for embedded devices," in Proc. IEEE Symposium Security Privacy, 2004: 272 – 282.

[24] A. Seshadri, M. Luk, A. Perrig, L. v. Doorn, and P. Khosla, "SCUBA: Secure code update by attestation in sensor networks," in Proc. 5^{th}ACM Workshop Wireless Security, 2006: 85 – 94.

[25] R. Sailer, X. Zhang, T. Jaeger, and L. v. Doorn, "Design and im – plementation of a TCG – based integrity measurement architecture," in Proc. 13th USENIX Security Symposium, vol. 13, IBM T. J. Watson Research Center, 2004.

[26] C. Krauss, F. Stumpf, and C. Eckert, "Detecting node compromise inhybrid wireless sensor networks Using attestation techniques," in Proc. Security and Privacy in Ad – hoc and Sensor Networks, vol. 4572/2007. Springer: Berlin/Heidelberg, 2007.

[27] H. Song, L. Xie, S. Zhu, and G. Cao, "Sensor node compromise detection: The location perspective," in Proc. International Conf. Wireless Commun. Mobile Computing, 2007: 242 – 247.

[28] Chen X, Makki K, Yen K, et al. Sensor network security: a survey [J]. Communications Surveys & Tutorials, IEEE, 2009, 11 (2): 52 – 73.

第 9 章　感知层嵌入式系统安全

物联网感知层设备很多都是基于嵌入式系统设计开发的。硬件平台和嵌入式操作系统是整个嵌入式系统的两个核心部分，硬件平台主要借助嵌入式微处理器来设计实现，操作系统通常则包括与硬件相关的设备驱动、系统内核、通信协议及图形界面等。物联网设备作为典型的嵌入式系统，在设计开发时同样需要注意安全问题。本章针对可信计算平台及典型的 TinyOS 嵌入式操作系统安全进行介绍。

9.1　平台安全——可信计算技术

9.1.1　可信计算技术概述

可信计算组织（Ttrusted Computing Group，TCG）对“可信”的定义是：“一个实体在实现给定目标时，若其行为总是如同预期，则该实体是可信的”（An entity can be trusted if it always behaves in the expected manner for the intended purpose）。这个定义将可信计算和当前的安全技术分开：可信强调行为结果可预期，但并不等于行为是安全的，这是两个不同的概念。根据英特尔的密码与信息安全专家大卫·格劳洛克（David Grawrock）的说法，如果你知道你的计算机中有病毒，这些病毒会在什么时候发作，了解会产生如何的后果，同时病毒也确实是这么运行的，那么这台计算机就是可信的。从 TCG 的定义来看，可信实际上还包含了容错计算里可靠性的概念。可靠性能保证硬件或软件系统的性能可预测。

可信计算为行为安全而生。据我国信息安全专家在《软件行为学》一书中描述，行为安全应该包括行为的机密性、行为的完整性、行为的真实性等特征。

从概念上来说，可信计算（Trusted Computing，TC）并非由可信计算组织（Trusted Computing Group，TCG，以前称为 TCPA）率先提出。可信这个概念早在彩虹系列的橘皮书中就有提及，它的目标就是提出一种能够超越预设安全规则，执行特殊行为的运行实体。操作系统中将这个实体运行的环境称为可信计算基（Trusted Computing Base，TCB）。

为了实现这个目标，人们从 20 世纪 70 年代之后就在做着不懈的努力。从应用程序层面、从操作系统层面、从硬件层面提出的 TCB 相当多。最为实用的是以硬件平台为基础的可信计算平台（Trustec Computing Platform，TCP），它包括安全协处理器、密码加速器、个人令牌、软件狗、可信平台模块（Trusted Platform Modules，TPM）及增强型 CPU、安全设备和多功能设备。

这些实例的目标是实现数据的真实性、数据的机密性、数据保护及代码的真实性、代码的机密性和代码的保护。

根据 S. W. Smith 前些年的著作《可信计算平台：设计与应用》（冯登国等翻译，清华大学出版社出版），这些平台的实现目的包括两个层面的意思：

① 保护指定的数据存储区，防止敌手实施特定类型的物理访问；

② 赋予所有在计算平台上执行的代码以证明它在一个未被篡改环境中运行的能力。

从广义的角度，可信计算平台为网络用户提供了一个更为宽广的安全环境，它从安全体系的角度来描述安全问题，确保用户的安全执行环境，突破被动防御打补丁方式。

9.1.2　TCG 可信计算平台体系结构及特征

1. 可信计算平台体系结构

当前计算机网络安全防范的重点主要是放在对服务器和网络的保护上，很少考虑网络终端接入者自身的安全，事实上大多数的攻击事件恰恰就是由终端不安全引发的。如果能够从网络终端系统平台建立起安全的防护体系，把不安全因素从源头控制好，就有望从根本上解决大多数的安全问题。可信计算平台主要就是利用这一思路，通过增强现有终端体系结构的安全性来保证整个系统的安全。其主要方法是在各种终端（包含 PC、手机及其他物联网终端等）设备上引入可信架构，通过其提供的安全特性来提高终端系统的安全性。终端可信的核心是称为可信平台模块（Trusted Platform Module，TPM）的可信芯片。

以 TPM 为基础的“可信计算”由用户的身份认证、平台完整性、应用程序的完整性、平台之间的可验证性等几个方面构成。图 9-1 描述了一个加入了可信计算模块的典型的 PC 终端框架，这个框架描述了 TPM 在 PC 平台中的位置。

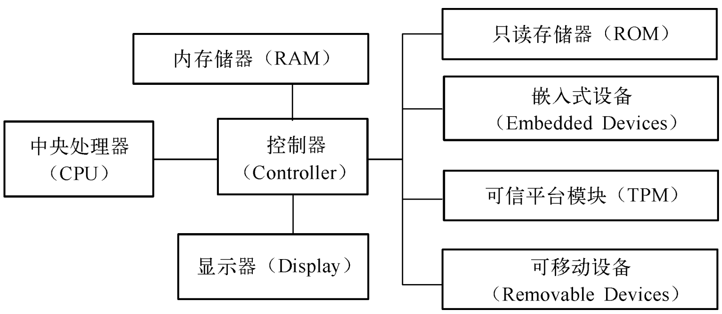


图 9-1　TCG 中的 PC 平台参考框架

可信平台以 TPM 为核心，但它并不仅仅由一块芯片构成，而是把 CPU、操作系统、应用软件、网络基础设备融为一体的完整体系结构。一个典型的可信平台体系结构如图 9-2 所示。

整个体系分为三层：TPM、TSS 和应用软件。

TPM 是可信平台的核心和基础，是由 TCG 定义的安全硬件子系统。TPM 软、硬件提供的密码功能有 RNG、哈希、HMAC、非对称密钥生成、非对称加解密等。它由 160 位的平台配置寄存器组成，PCR 用来保存 SHA-1 操作的结果。TPM 至少拥有 8 个 PCR，用来在可信启动过程中记录软件完整性度量的结果。

TSS 是软件协议栈，处在 TPM 之上，应用软件之下，称为可信软件栈，它提供了应用程序访问 TPM 的接口，同时进行对 TPM 的管理。图 9-2 中的虚线中间部分为 TSS 协议栈结构，TSS 分为四层：工作在用户态的 TSP（Trusted Service Provider）、TCS（TSS Core Services）、TDDL（TPM Device Driver Library）和内核态的 TDD（TPM Device Driver）。它允许应

用程序和可信移动设备 TMD（Trusted Mobile Device）中的可信平台模块（TPM）进行通信。它为 TPM 功能的使用提供了接口，如鉴权（Authentication）、授权、保护存储和认证（Attestation）。它还为利用 TPM 基于硬件的密码服务提供了接口，允许应用程序使用 TPM 来生成密钥、加密/解密、签名等。TSS 为应用程序利用 TPM 服务提供了同步访问机制。除此之外，TSS 还管理 TPM 资源。

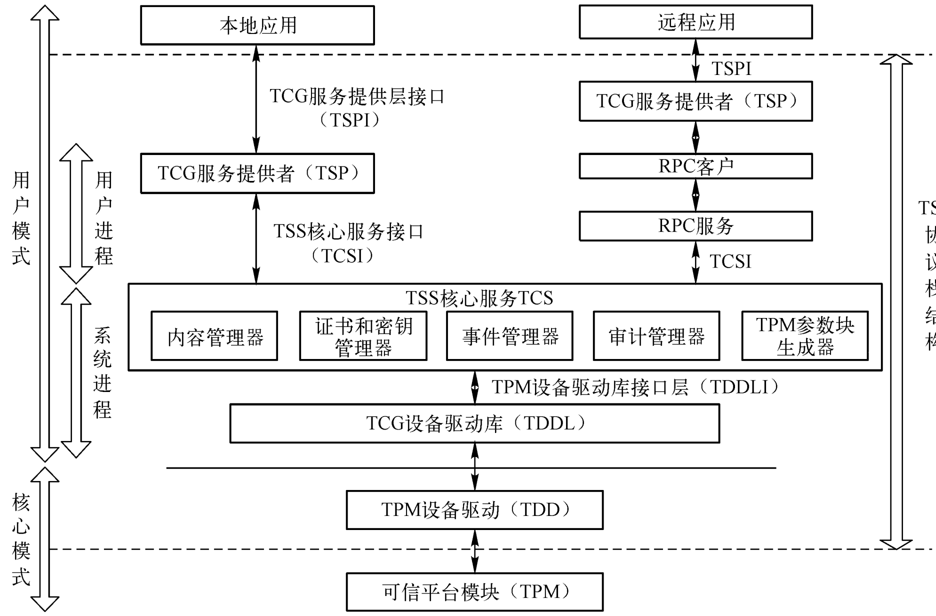


图 9-2　TCG 可信平台体系结构

TCG 软件栈定义了两种模式——用户模式和内核模式，三个接口——TCG 服务提供者接口（TSPI）、TSS 核心服务接口（TCSI）和 TPM 设备驱动接口（TDDLI）。内核模式包括 TPM 和 TPM 的设备驱动，用户模式又分为用户进程和系统进程，用户进程指的是 TCG 服务提供者提供给外部应用程序的服务操作，系统进程指的是 TCG 设备驱动库提供的 TSS 核心服务操作。三个接口与不同层次的计算平台服务通信，其中，TDDL 接口提供了用户模式和内核模式之间的过渡，它工作在 TPM 设备和设备驱动层之上；TCS 为一组普通的平台服务提供接口，它保证了即使在一个平台上可能存在多个 TCG 服务提供者，它们都只展示相同的行为；TSP 为 TPM 提供了一个 C 接口，它作为应用程序驻扎在公共的进程地址空间。TSP 提供两种服务：上下文管理和密码功能。上下文管理考虑到应用程序和 TSP 资源的使用效率，并为此提供动态的操作。密码功能的使用则是为了充分地利用 TPM 受保护的功能。

1）TCG 服务提供层（TSP）

TSP 提供应用程序访问 TPM 的 C + + 界面，基于一个面向对象的底层结构，驻留在与应用程序一样的进程地址空间（都是用户进程）。授权协议在这一层通过一个在这层编码的用户接口，也可通过 TCS 层的回调机制（如果调用者是远程）来实现。

TSP 提供两种服务：上下文（context）管理和密码功能。上下文管理器产生动态句柄，以便高效地使用应用程序和 TSP 资源。每个句柄提供一组相关 TCG 操作的上下文。应用程

序中不同的线程可能共享一个上下文，也可能每个线程获得单独的上下文。为了充分利用TPM 的安全功能，这层也提供了密码功能。但是内部数据加密对接口是保密的，如报文摘要和比特流的产生功能等。

2）TCG 核心服务层（TCS）

TCS 提供一组标准平台服务的 API（Application Programming Interfaces）接口。一个 TCS 可以提供服务给多个 TSP。如果多个 TCG 服务提供者都基于同一个平台，TCS 保证它们都将得到相同的服务。TCS 提供了 4 个核心服务：

① 上下文管理，实现到 TPM 的线程访问；

② 证书和密钥的管理，存储与平台相关的证书和密钥；

③ 度量事件管理，管理事件日志的写入和相应 PCR（Platform Configuration Registers，平台配置寄存器，这种寄存器位于 TPM 内部，仅仅用来装载对模块的度量值，大小为160bits）；

④ 参数块的产生，负责对 TPM 命令序列化、同步和处理。

3）TCG 设备驱动库（TDDL）

TDDL 是用户态和内核态的过渡，仅仅是一个接口而已。它不对线程与 TPM 的交互（Interaction）进行管理，也不对 TPM 命令进行序列化（Serialization）。这些在高层的软件堆完成。由于 TPM 不是多线程的，一个平台只有一个 TDDL 实例（Instance），从而只允许单线程访问 TPM。TDDL 提供开放接口，使不同各厂商可以各自自由实现 TDD 和 TPM。

2. 可信计算平台特征

一个可信平台要达到可信的目标，最基本的原则就是必须真实报告系统的状态，同时决不暴露密钥和尽量不表露自己的身份。这就需要三个必要的基本特征：保护能力（Protected Capabilities）；证明（Attestation）；完整性度量、存储和报告（Integrity Measurement，Storage and Reporting）。

1）保护能力

保护能力是唯一被许可访问保护区域（Shielded Locations）的一组命令，而保护区域是能够安全操作敏感数据的地方（如内存、寄存器等）。TPM 通过实现保护能力和被保护区域来保护和报告完整性度量。除此之外，TPM 保护能力还有许多的安全和管理功能，如密钥管理、随机数生成、将系统状态值封印（Seal）到数据等。这些功能使得系统的状态任何时候都可知，同时可以将系统的状态与数据绑定起来。由于 TPM 的物理防篡改性，这也就起到了保护系统敏感数据的功能。

2）证明

证明是确认信息正确性的过程。通过这个过程，外部实体可以确认保护区域、保护能力和信任源，而本地调用则不需要证明。通过证明，可以完成网络通信中身份的认证，而且由于引入了 PCR 值，在身份认证的同时还鉴别了通信对象的平台环境配置，这大大提高了通信的安全性。

证明可以在不同层次进行：基于 TPM 的证明是一个提供 TPM 数据的校验的操作，这是通过使用 AIK（Attestation Identity Key）对 TPM 内部某个 PCR 值的数字签名来完成的，

AIK 是通过唯一秘密私钥 EK（Endorsement Key，签注密钥）获得的，可以唯一地确认身份；针对平台（To the platform）的证明则通过使用平台相关的证书或这些证书的子集来提供证据，证明平台可以被信任以做出完整性度量报告；基于平台（of the platform）的证明通过在 TPM 中使用 AIK 对涉及平台环境状态的 PCR 值进行数字签名，从而提供了平台完整性度量的证据。

3）完整性度量、存储和报告

完整性的度量是一个过程，包括：获得一个关于平台的影响可信度特征的值（metrics），存储这些值，然后将这些值的摘要放入 PCRs 中。通过计算某个模块的摘要与期望值的比较，就可以维护这个模块的完整性。在 TCG 的体系中，所有模块（软件和硬件）都被纳入保护范围内，假如有任何模块被恶意感染，它的摘要值必然会发生改变，使我们可以知道它出现了问题，虽然还不能知道问题是什么。通过这种方式，就可以保护所有已经建立 PCR 保护的模块。

另外，平台 BIOS 及所有启动和操作系统模块的摘要值都将存入特定的 PCR，在进行网络通信时，可以通过对通行方 PCRs 值的校验确定对方系统是否可信（即是否感染了病毒、是否有木马、是否使用盗版软件等）。

度量必须有一个起点，这个起点必须是绝对可信的，叫作度量可信根（Root of Trust for Measurement，RTM）。一次度量叫作一个度量事件（Event），每个度量事件由以下两类数据组成。

① 被度量的值：嵌入式数据或程序代码的特征值（Representation）。

② 度量摘要：这些值的散列。

完整性报告用来证明完整性存储的过程，展示保护区域中完整性度量值的存储，依靠可信平台的鉴定能力证明存储值的正确性。TPM 本身并不知道什么是正确的值，它只是忠实地计算并把结果报告出来。这个值是否正确还需要执行度量的程序本身通过度量存储日志（Stored Measurement Log，SML）来确定。此时的完整性报告使用 AIK 签名，以鉴别 PCR 的值。按照“可信”的定义，完整性度量、存储、报告的基本原则就是：许可平台进入任何可能状态（包括不期望的或不安全的），但是不允许平台提供虚假的状态。

除了计算的散列值存在 PCR 里外，还需要存储期望值。SML 保存着有关系（Related）的被度量值的序列，每个序列公用一个通用摘要。这些被度量的值附加在通用摘要之后被再次散列，通常称之为摘要的扩展。扩展保证了不会忽视这些有关系的被度量值，同时可以保证操作的顺序。SML 可能会非常大，需要存在硬盘上，不过由于都是散列值，所以不需要 TPM 提供保护。完整性报告协议如图 9-3 所示。

完整性报告协议描述了对一个事件的度量进行校验的完整性报告协议的执行过程：

- 一个远程的外界访问者（Challenger）向 TCS 发送请求，需要一个或多个 PCR 值；
- TCS 读取 SML 以获得度量时间的数据；
- TCS 发送命令到 TPM，请求获得 PCR 值；
- TPM 使用 AIK 对 PCR 值签名；
- TCS 从知识库里收集用来证明 TPM 平台的证书（AIK 证书、平台证书等）；
- 访问者在本地校验请求，如果校验不通过，则说明存在问题，但无法获得任何关于错误的信息。

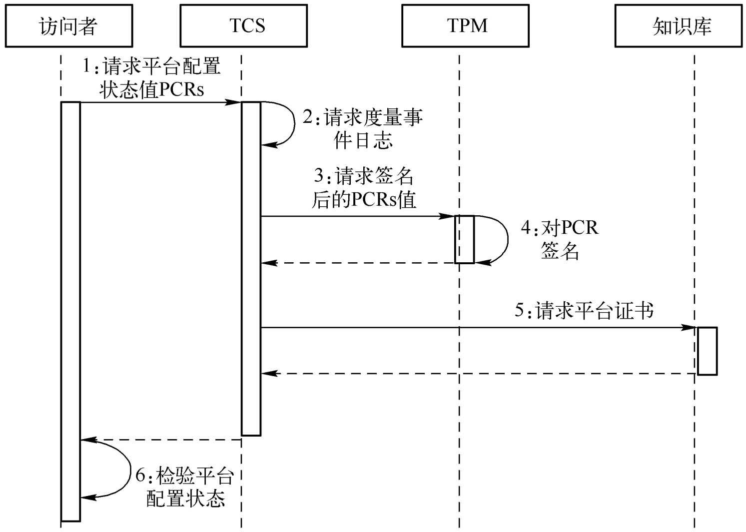


图 9-3　完整性报告协议

9.1.3　TPM 可信平台模块

可信计算平台基于可信平台模块（TPM），以密码技术为支持、安全操作系统为核心。安全操作系统是可信计算平台的核心和基础，没有安全的操作系统，就没有安全的应用，也不能使 TPM 发挥应有的作用。

在可信计算技术体系中，最核心的就是 TPM 芯片。可信平台模块可以看成一个完整的计算机，有处理器、协处理器、存储单元和操作系统等。TPM 至少需要具备四个主要功能：对称/非对称加密、安全存储、完整性度量和签名认证。数据的非对称加密和签名认证是通过 RSA 算法来实现的，而完整性度量则通过高效的 SHA－1 散列算法来完成，对称加密可以使用任意算法，既可以使用专用协处理器也可以使用软件来完成。图 9-4 是 TCG 可信平台模块结构图。

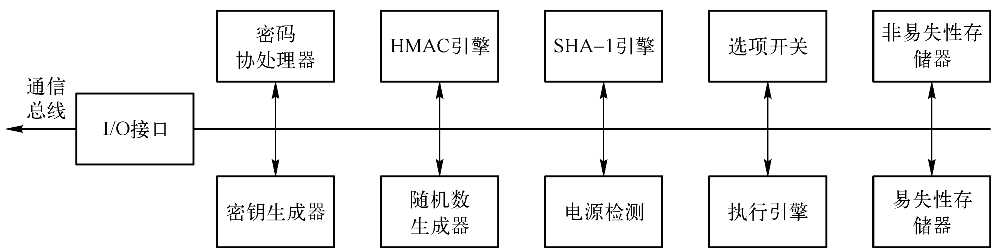


图 9-4　TCG 可信平台模块结构

I/O 接口：TPM 的输入/输出接口，负责管理通信总线上的信息流，主要任务包括执行内部总线和外部总线之间转换的通信协议，并向合适的组件发送消息，执行对 TPM 进行操作的安全策略。

密码协处理器：密码协处理器组件负责 RSA 运算的实现，它内含一个执行运算的 RSA 引擎，提供对内外的数字签名、内部存储和传输数据的加解密功能，以及密钥的产生、安全

存储和使用等管理功能。

密钥生成器：密钥生成器组件负责创建 RSA 密钥对和对称密钥。TCG 没有限制非对称密钥或对称密钥的生成次数。

HMAC 引擎：HMAC 引擎组件负责确认报文数据是否以正确的方式为 TPM 提供信息，它可以发现数据和命令错误或被篡改的情况。

随机数生成器：随机数生成器组件是 TPM 中随机数的产生来源，负责产生各种运算所需要的随机数。TPM 利用这些随机数值来生成现时、对称密钥和签名中的随机数。

SHA－1 引擎：SHA－1 引擎组件通过运行可靠的杂凑算法执行杂凑操作。TPM 向外部提供杂凑接口以支持在平台导入阶段进行度量，并允许具有有限能力的环境访问杂凑函数。

电源检测：电源检测组件管理着 TPM 的电源状态，帮助 TPM 在电源状态发生变化时采取适当的限制措施。

选项开关：选项开关组件提供对 TPM 进行功能开启/关闭、使能/失能和激活/去激活的机制，通过改变一些永久性的可变标志位，可以设置 TPM 的功能选项。

执行引擎：执行引擎组件负责执行经过 I/O 接口传送给 TPM 的命令，它是一个保证操作被适当隔离和保护区域被保护的关键组件。

非易失性存储器：非易失性存储器组件用来保存永久身份和与 TPM 相关联的状态。

易失性存储器：易失性存储器组件用来保存 TPM 运行时的临时数据。

以上若干组件构成一个有机、统一的安全执行环境，作为嵌入式的芯片部件，它们高度集成，并且功能完善。密钥和授权信息处于底层的 TPM 所提供的硬件加密保护之下，攻击者只有攻破 TPM 才能攻破系统防护，这样 TPM 成为系统可信的最低层次，是系统可信的基础。

在集成有 TPM 芯片的计算平台上，用户通过 TPM 提供的 API 接口编程函数来使用 TPM。用户除了通过 TCG 制定的 TPM 规范中所声明的 API 访问 TPM 外，不能够通过其他的软件方法访问 TPM。TPM 芯片实现 TPM 时，把 TPM 制造成具有防软件篡改和一定的硬件篡改功能的芯片。保证 TPM 只能通过 TCG 规范制定的 API 访问，是 TPM 完成其功能的基础。

9.2　平台安全——TrustZone 技术

9.2.1　TrustZone 技术概述

对于可信硬件平台，ARM 公司给出了另外一套解决方案，那就是 TrustZone 技术。ARM TrustZone 技术的目的在于使一个嵌入式设备同时受益于一个功能丰富的开放操作环境和一个健壮的安全解决方案，无论操作环境有多么不可信，设计良好的系统硬件架构和合适的安全软件设计依然能确保敏感数据的安全性。

TrustZone 硬件架构的目的在于提供一个安全的框架，在该框架下，嵌入式设备能够抵御它将遇到的许多的特定攻击。TrustZone 技术并不是提供一个固定的且能针对所有安全隐患的解决方案，而是提供了一系列的基础设施，片上系统（SoC）的设计者可以从中选择出自己所需要的组件，并在安全环境中实现他们需要的一些特定的功能。TrustZone 硬件架构

的主要目标实际上是非常简单的，即使一个可编程环境下的几乎所有资产的机密性和完整性都能受到保护，从而抵御住特定的攻击。如果一个开放平台拥有上述特征，就可以将其构建成一套广泛适用的安全解决方案，并且该方案与传统方法相比较，成本效益更好。

为了实现系统的安全性，TrustZone 技术将 SoC 的硬件资源和软件资源进行了划分，并将其隔离为两个执行环境，即安全环境（Secure world）和普通环境（Normal world），安全环境执行安全子系统中的任务，普通环境则处理其他所有事务。在具有 TrustZone 技术的设备的硬件逻辑中，AMBA3 AXI 总线结构在两个执行环境中间建立了一个强大的安全边界，使得安全环境中的任意资源都不能被普通环境的组件所访问。然后将敏感数据存放在安全环境中及在安全环境下的处理器上运行软件程序，这样就能够保护几乎所有的有价值信息免于可能会受到的许多攻击。如今，TrustZone 技术的安全扩展机制已经在一些 ARM 处理器内核中实现了。支持 TrustZone 技术的 ARM 处理器有：

- ARM Cortex－A15；
- ARM Cortex－A9；
- ARM Cortex－A8；
- ARM Cortex－A7；
- ARM Cortex－A5；
- ARM1176。

这一功能的增加将使得单个的处理器内核能够安全有效地同时执行来自安全环境和普通环境的代码，这样就不需要给安全环境提供一个专用的安全处理器，将极大地减少芯片面积和功耗，这也使得高性能的安全软件与普通环境的操作系统能同时运行。最后，TrustZone 硬件架构还是一个有安全意识的调试基础设施，它可以控制对安全环境的调试访问而不影响普通环境调试的可见性。

9.2.2　TrustZone 硬件架构

1. AMBA3 AXI 系统总线

扩展总线设计中最重要的一个功能是为每个主要系统总线的读写通道都添加一个额外的控制信号，这些比特位被称为 NS 比特位（Non－Secure），其在 AMBA3 AXI 总线协议规范中被定义，读写控制信号如下。

AWPROT[1]：总线写事务控制信号，低电平为安全环境，高电平为非安全环境（普通环境）。

ARPROT[1]：总线读事务控制信号，低电平为安全环境，高电平为非安全环境。

所有的总线主设备在发起一个新的事务时将会对 NS 比特位进行设置，系统总线或它的从设备解码逻辑必须能够解释这些信号位，以确保安全分离机制不会被违背。例如，所有非安全环境中的主设备必须将硬件中的 NS 位设置为高电平，确保它们不能对安全环境的从设备进行访问，地址解码器中的访问地址也不能与任意的安全从设备进行匹配。

2. AMBA3 APB 外围总线

TrustZone 架构最有用的特性之一就是能够保护外围设备，如中断控制器、计时器和用

户 I/O 设备。这个特性使得安全环境得到扩展，可以解决更广泛的安全问题，而不仅仅是提供一个安全的数据处理环境。在 AMBA3 规范中介绍了一个低门数、低带宽的外围设备总线即 APB 总线，它通过一个 AXI – to – APB 桥与系统总线相连接。APB 总线并没有与 NS 比特位对应的信号位，这使得现有的 AMBA2 APB 外围设备能与 TrustZone 技术相兼容。这样，负责 APB 外围设备安全的责任就交给了 AXI – to – APB 桥。该桥的主要任务是能够拒绝事务中不适合的安全设置及不得将这些请求转发到外围设备去。

3. 内存别名

TrustZone 技术不仅在总线事务中添加了一个 NS 位，还同时在系统的 cache tags 中添加了相应位，可以将这个添加的比特位视为地址空间的第 33 位，这样安全环境和非安全环境就可以分别对应一个 32 位的物理地址空间。假设一个安全环境的主设备想要访问非安全环境的从设备，因为安全环境和非安全环境分别对应不同的物理地址空间，那么如何解决这个问题？TrustZone 技术的硬件架构提供了内存别名功能，在这个内存别名系统中，同一个内存位置在地址映射中显示为两个完全不同的地址，其中一个地址用于安全环境访问，另一个就用于非安全环境访问。但是这种别名使用可能引起数据一致性问题，有可能在缓存中会同时存在多个值代表相同的数据，因此系统设计者必须意识到潜在的数据一致性问题，并采取措施来避免它们。

4. 环境切换

在 TrustZone 技术中，每个处理器内核将提供两个虚拟内核，一个被认为是安全的，另一个是不安全的，在这两个环境之间进行上下文切换的机制被称为监视器模式，如图 9–5 所示。两个虚拟内核通过时间片方式进行轮转。

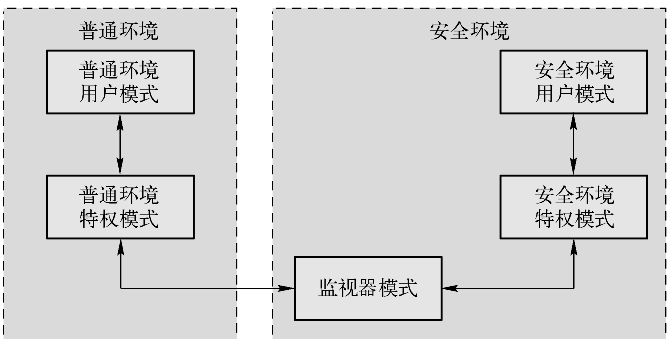


图 9–5　TrustZone 安全扩展模型

从普通环境进入监视器模式是被严格控制的，这个机制被视为监视器模式软件异常。想要进入监视器模式，可以通过执行一条专门的指令，即 Secure Monitor Call（SMC）指令来触发，或者通过硬件异常机制，如对 IRQ、FIQ、外部数据终止等进行配置，使处理器能够切换到监视器模式。

在监视器模式下运行的软件已被定义，它的主要功能就是保持当前环境的状态和恢复即将切换过去的环境的状态，然后执行一个从异常状态返回的操作来重启所恢复的环境中的处理器。处理器运行的操作环境可以通过系统控制协处理器 CP15 中的 Secure Configuration

Register（SCR）中的 NS 比特位来表明。但在监控模式下，SCR 中 NS 位的值不论是什么，处理器都运行在安全环境中，只是在访问分组寄存器时，若 SCR 中的 NS 位被设置为 1（即普通环境），则将访问普通环境中的备份数据。

5. 内存系统安全

TrustZone 技术下，内存被分离成了两个环境（安全的和非安全），同理，处理器内核的一级缓存（L1）中的用于存储数据的组件也被分离成两个环境。L1 内存系统中的主要部件是 Memory Management Unit（MMU），它的作用是将虚拟地址映射为物理地址，地址的转换操作由一个软件控制转换表来进行管理。在一个有 MMU 但是不含安全扩展的 ARM 内核中，只存在一个虚拟地址到物理地址的映射。若有 TrustZone 安全技术扩展，硬件将会提供两个虚拟 MMU，分别对应两个虚拟处理器，这样每个执行环境都能够拥有一个本地的地址转换表，能够独立控制虚拟地址到物理地址的映射。若缓存不仅能存储数据，还能存储数据的安全状态，将消除环境切换过程中的缓存刷新，也可使高性能软件能够在两个环境边界之间进行通信。为了实现具有这样功能的缓存，L1、L2 和更多级的缓存都被进行了扩展，为它们添加了一个标签位，用于记录安全状态。

6. 安全中断

在 ARM TrustZone 安全机制扩展模型中，IRQ 常常被用作普通环境的中断源，而安全环境的中断源往往选择 FIQ，这是因为在大多数操作系统环境中，IRQ 是使用最多、最常见的中断源，用 FIQ 作为安全中断源将最小限度地修改已有软件。执行监控器代码时通常应该关闭中断。

7. 安全处理器配置

为了使两个虚拟 CPU 的代码能够独立地运行，硬件对 CP15 上的配置选项进行了严格的管理。敏感的配置选项和全局资源只能在安全环境中进行写操作，但一般可以在普通环境进行读操作。只在本地环境中设置，通常在硬件中会进行分组，使每个环境能独立地控制配置选项。

9.2.3 TrustZone 软件架构

在 Soc 硬件上要实现安全执行环境还需要一些安全软件运行在其上并使用安全环境中存放的敏感资源。ARM TrustZone 安全扩展是一个开放的组件，任意的开发者都可以创建一个安全软件来满足客户的需求。

1. 软件架构

安全环境所能处理的资源对软件架构的整体结构有很大的影响，在具有 TrustZone 功能的处理器内核上可以实现多种软件架构。最复杂的方法是在安全环境运行一个专用的安全操作系统，最简单的方法是在安全环境放置一个同步代码库。在这两个方法之间还有许多中间选项。如图 9-6 所示为一种安全执行环境可能的软件架构。

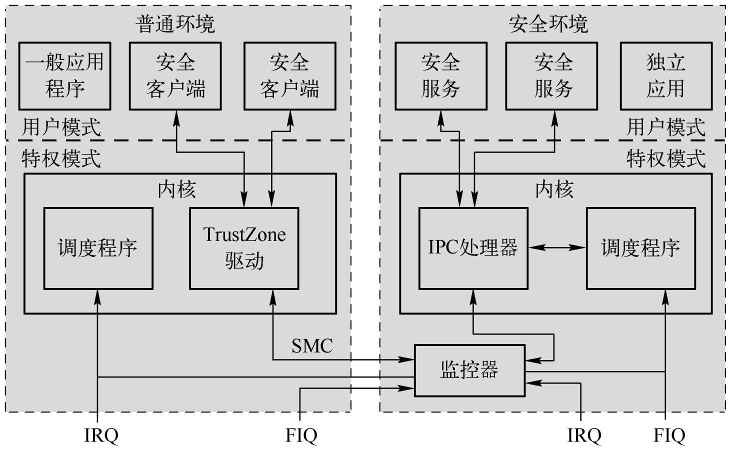


图 9-6　一种基于 TrustZone 的软件架构

在安全环境运行一个专用操作系统的方法，虽然实现起来很复杂，但系统功能强大。它可以模拟并发执行多个安全环境的应用、下载新的安全应用，以及将安全环境的任务完全独立于普通环境等。然而许多用例是不需要这么复杂的安全操作系统的。对许多应用程序来说，在安全环境中，一个简单的代码库就足够处理一个任务了。这个代码库是通过软件调用由普通环境的操作系统来全权计划和管理的。如此，安全环境相当于是普通环境的一个子系统，它不能独立运行，但是降低了复杂度，便于开发实现。位于这两个极端方法之间，还有一系列的选择。例如，可以设计一个具有多任务的操作系统，但是它没有专用中断源，而是通过普通环境来提供。这种设计在普通环境停止提供虚拟中断时容易受到拒绝服务攻击，但是对许多任务来说这种攻击不会造成严重问题。

2. 安全系统的引导

在一个安全系统的生命周期内，最重要的时刻是引导时期。因为有许多攻击者会选择在设备断电时对软件进行攻击，如篡改存放在 Flash 中的安全系统软件。当系统启动时，如果直接从 Flash 去引导镜像而不对镜像进行验证，这个系统很有可能是不安全的。在 TrustZone 扩展机制中，原则之一是认为在安全环境中从信任根生成的可信链是安全的，是不容易被篡改的。因此，安全启动的顺序如图 9-7 所示，安全引导在每个安全环境的引导过程中都增加了一个加密检查操作，维护了在安全环境运行的软件镜像的完整性，防止任何未经授权的或恶意修改的软件被运行。安全环境在上电后先启动，可以保证在普通环境修改系统之前，任意的安全问题都能够先被检查。

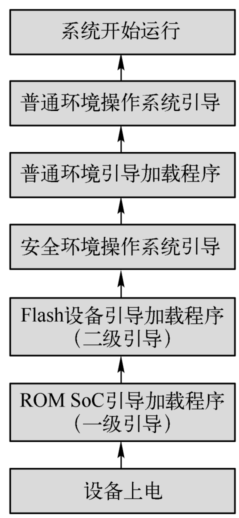


图 9-7　TrustZone 处理器中的系统引导加载顺序

ARM TrustZone 是一种软硬件结合的安全解决方案，由硬件来实现安全环境与普通环境的隔离，软件提供基本的安全服务和接口。Global Platform 国际标准组织制定了一套支持 ARM 隔离技术的软硬件技术规范，并将隔离出来的安全环境称为可信执行环境（Trusted Execution Environment，TEE）。与 TEE 相对

应的是 Rich Execution Environment（REE），REE 为应用程序提供了一个资源丰富的通用操作系统，普通的应用程序便是运行在其上的。TEE 是一个独立的执行环境，它为 REE 和在其内部运行的应用程序提供了多种安全服务，TEE 主要通过一组 API 接口来实现 REE 与 TEE 中可信应用软件的通信。一个可信执行环境应满足以下几点：

① 在可信执行环境中运行的任意代码在真实性和完整性上是可信的；

② 除代码外的其他资产也应受到机密性保护，可信执行环境（ TEE）应能够抵挡所有已知的软件和远程攻击，以及一组外部硬件攻击；

③ 通过调试和测试功能，资产和代码都能免受未经授权的跟踪和控制。

ARM 现已将其 TrustZone API 提供给 Global Platform，该 API 已发展为 TEE 客户端 API。ARM 还与其他一流公司合作开发在可信 OS 与可信应用程序之间进行交互的 TEE 内部 API。预计 TEE 的标准化将会促使可信应用程序部署的快速增长。

9.3 TinyOS 操作系统及其安全技术

9.3.1 TinyOS 操作系统概述

安全的物联网设备，不但硬件平台要可信，其运行的操作系统也要安全。TinyOS 是由加州大学伯克利分校开发的一个开源的嵌入式操作系统，采用 nesC 语言编写，主要应用于无线传感器网络方面。和传统意义上的操作系统不同，TinyOS 是一个适用于网络化嵌入式系统的编程框架，通过在这个框架里将用户设计的一些组件和操作系统的必要组件连接起来，就能方便地编译出面向特定应用的操作系统，这对于硬件资源极为有限的系统来说非常重要。目前在世界范围内，有超过 500 个研究小组或公司在使用这个微型的操作系统。

1. TinyOS 系统设计思路

由于传感器网络的特殊性，需要操作系统能够高效地使用传感器节点的有限内存、低功耗处理器、有限的电源、低速通信设备、多样的传感器，且能够对各种特定应用提供最大的支持。在面向传感器网络的操作系统支持下，多个应用（如计算、存储和通信等）可以并发地使用系统资源。针对这些要求，研究人员在设计和实现 TinyOS 操作系统时，提出了以下几个必须遵循的设计要求。

① 能在有限的资源上运行：要求执行模式允许在单一的协议栈上运行。

② 允许高度的并发性：要求执行模式能对事件做出快速的直接响应。

③ 适应硬件升级：要求组件和执行模式能够应对硬件/软件的替换。

④ 支持多样化的应用程序：要求能够根据实际需要，裁减操作系统的服务。

⑤ 鲁棒性强：要求通过组件间有限的交互渠道，就能应对各种复杂情况。

⑥ 支持一系列平台：要求操作系统的服务具有可移植性。

为此，在 TinyOS 系统的设计之初，加州大学的研究人员确定了全组件化、事件驱动、无内核和用户空间区分的设计原则来满足无线传感器网络的特殊需求。应用程序根据需要选配、修改和创建组件，使系统开销最小化。组件与组件之间通过“命令”（Command）和“事件”（Event）相联系，“命令”向下调用（call down），“事件”则向上调用（call up）。

因此，TinyOS 被称为“事件驱动”的操作系统。

到目前为止，无线传感器网络还没有一个公认的体系结构，不像传统网络那样有明显的层次划分和明确的功能抽象与定义。在 TinyOS 1. x 中，模糊了这些概念，采用了平面型的设计思想。但研究和应用出现的问题启发人们重新审视无线传感器网络操作系统的设计原则。例如，无线传感器网络要不要层次性的结构，如何抽象和定义各层的功能，如何将无线传感器网络的特殊需要（如分组聚合、数据网内处理、数据查询、分布式存储和时间同步的 MAC 层增强机制等）抽象定义在恰当的功能层中，IP 网络中的地址方案和基于此的路由协议是否适合传感器网络等。对于这些基本问题，有些已经有了倾向性的阶段性结论，有些依旧需要进一步研究和讨论。在 TinyOS 2. 0 中，从一些增强的功能和新的特性上可以看出这几年传感器网络研究的新进展。例如，在 TinyOS 1. x 中，因为没有过多考虑异构硬件平台的支持，移植工作难度较大；而 TinyOS 2. 0 定义了 3 个层次的硬件抽象结构，提供了独立的硬件边界，强化了对异构平台的支持，同时保持了针对特定硬件进行优化设计的灵活性。为了更好地支持应用开发，TinyOS 2. 0 还提供了更加丰富的业务库，允许在最基本的通信机制和最基本的组件（如定时器等）之上进行高层抽象，提高代码的可重用性。TinyOS 2. 0 较 1. x 有了本质的变化，但它离理想的无线传感器网络操作系统仍有距离，因为有许多问题依旧没有解决。人们仍然没有找到合理的方案，还不是完全清楚该如何设计。因此，还需要进行大量的研究工作，为无线传感器网络操作系统的设计实现寻找更充分的支撑依据。

2. TinyOS 体系结构

TinyOS 操作系统最初是通过汇编语言和 C 语言编写的。但 C 语言不能有效、方便地支持面向无线传感器网络的应用程序和操作系统的开发，且目标代码比较长。为此，科研人员经过研究，对 C 语言进行了一定扩展，提出了支持组件化编程的 nesC（C language for network embedded systems）语言，其最大的特点是把组件化/模块化思想和基于事件驱动的执行模式结合起来。TinyOS 操作系统本身和基于 TinyOS 的应用程序基本上都采用 nesC 语言编写，这提高了应用开发的方便性和应用执行的可靠性。即使在只有少量 ROM 的情况下，TinyOS 操作系统也能支持高度的并发处理及复杂的协议和算法，而且能高效地运行在无线传感器网络环境中。

TinyOS 操作系统采用了组件的结构，它是一个基于事件的系统。系统本身提供了一系列的组件供用户调用，其中包括主组件、应用组件、执行组件、感知组件、通信组件和硬件抽象组件，其层次结构如图 9–8 所示。组件由下到上通常可以分为 3 类：硬件抽象组件、综合硬件组件和高层软件组件。硬件抽象组件将物理硬件映射到 TinyOS 的组件模型；综合硬件组件模拟高级的硬件行为，如感知组件、通信组件等；高层软件组件实现控制、路由及数据传输等应用层的功能。高层组件向底层组件发出命令，底层组件向高层组件报告事件。TinyOS 的层次结构就如同一个网络协议栈，底层的组件负责接收和发送最原始的数据位，而高层的组件对这些数据进行编码、解码，更高层的组件则负责数据打包、路由选择及数据传输。

调度器具有两层结构：第 1 层维护着命令和事件，它主要是在硬件中断发生时对组件的状态进行处理；第 2 层维护着任务（负责各种计算），只有当组件状态维护工作完成后，任务才能被调度。由前所述，TinyOS 调度模型主要有以下几个特点。

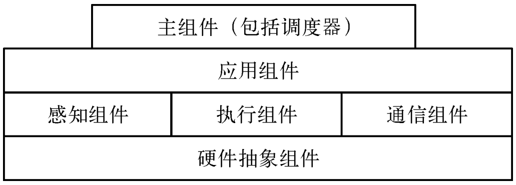


图 9-8　TinyOS 体系结构

① 任务单线程运行到结束，只分配单个任务栈，这对内存受限的系统很有利。

② 没有进程管理的概念，对任务按简单的 FIFO 队列进行调度。

③ FIFO 的任务调度策略具有能耗敏感性。当任务队列为空时，处理器进入休眠，随后由外部事件唤醒 CPU 进行任务调度。

④ 两级的调度结构可以实现优先执行少量与事件相关的处理，同时打断长时间运行的任务。

⑤ 基于事件的调度策略，只需要少量空间就可获得并发性，并允许独立的组件共享单个执行上下文。与事件相关的任务可以很快被处理，不允许阻塞，具有高度并发性。

⑥ 任务之间互相平等，没有优先级的概念。

在 TinyOS 程序模型中，处于最上层的是主组件，即 Main 组件。该组件由操作系统提供，节点上电后会首先执行该组件中的函数，其主要功能是初始化硬件、启动任务调度器及执行应用组件的初始化函数。每个 TinyOS 程序应具有至少一个应用组件，即用户组件。该应用组件通过接口调用下层组件提供的服务，实现程序的逻辑功能，如数据采集、数据处理或数据收发等。因此，应用组件的开发是 TinyOS 程序设计的重点。

一个完整的应用系统由一个内核调度器（简称调度器）和许多功能独立且相互联系的组件构成，应用程序与组件一起编译成系统。因此，可以把 TinyOS 系统和在其上运行的应用程序看成一个大的“执行程序”。现有的 TinyOS 系统提供了大多数传感网硬件平台和应用领域里都可用到的组件，如定时器组件、传感器组件、消息收发组件及电源管理组件等，从而把用户和底层硬件隔离开来。在此基础上，用户只需要开发针对特殊硬件和特殊应用需求的少量组件，大大提高了应用的开发效率。

3. 技术特点

TinyOS 操作系统本身在软件结构上就体现了一些已有的研究成果，如组件化编程方式（component-based programming）、事件驱动（event driven）模式、轻量级线程（lightweight thread）技术、主动消息（active message）通信技术等。这些研究成果最初并不是用于面向传感器网络的操作系统的，例如，轻量级线程和主动消息主要用于并行计算中的高性能通信。但经过对传感器网络应用系统的深入研究后发现，上述技术有助于提高传感器网络的性能，在发挥硬件功能的同时能降低其功耗，并且简化了应用程序的开发。

TinyOS 系统的技术优势主要体现在以下四个方面。

1）组件化编程

无线传感器网络既具有多样化的上层应用，又强调系统的节能性要求。为此，TinyOS 系统采用一种基于组件的体系结构，这种体系结构已经被广泛应用在嵌入式操作系统中。组

件就是对软、硬件进行功能抽象。整个系统由组件构成，通过组件提高软件重用度和兼容性，程序员只需要关心组件的功能接口和自己的业务逻辑，而不必关心组件的具体实现，从而能够提高编程效率，快速实现各种应用。

同时，TinyOS 程序采用的是模块化设计，只包含必要的组件，提高了操作系统的紧凑性。这样既便于上层应用的开发，也有利于程序的快速执行。这样设计的程序内核往往都很小，其内核代码和数据大概在 400B 左右，能够突破传感器存储资源少的限制，使得 TinyOS 系统可以有效地运行在无线传感器网络节点上，并负责执行相应的管理工作。

2）事件驱动机制

针对无线传感器网络内节点众多，以及并发操作频繁的工作方式，TinyOS 采用了事件驱动的运行机制。TinyOS 的应用程序都基于事件驱动模式，事件处理完成后处于长期的低功耗状态，只有事件来临时才会触发来唤醒传感器工作。事件相当于不同组件之间传递状态信息的信号。当事件对应的硬件中断发生时，系统能够快速地调用相关的事件处理程序，迅速响应外部事件，并且执行相应的操作任务。因此，事件可称为中断处理线程，常用于时间要求很严格的应用中。

TinyOS 中程序的运行由一个个事件驱动。数据包收发、传感器采样等操作引发的硬件中断会触发底层组件中的事件处理程序，对该中断做初步处理后再触发上层组件的事件，通知上层组件对该事件做进一步处理。事件驱动机制可以使 CPU 在事件产生时迅速执行相关任务，并在处理完成后进入休眠状态，有效地提高了 CPU 的使用率，并达到节能的目的。

3）轻量级线程技术及两层调度方式

TinyOS 提供任务和硬件事件处理两级调度体系。轻量级线程，即任务，用在对于时间要求不是很高的应用中。任务之间是平等的，不能相互抢占，按先入先出队列（First Input First Output，FIFO）进行调度。轻线程是针对节点并发操作可能比较频繁，且线程比较短的问题提出的。由于传感器节点的硬件资源有限，短流程的并发任务可能频繁执行，传统的进程或线程调度算法会在无效的进程切换过程中产生大量能耗，故无法应用于传感器网络的操作系统。轻量级线程技术和基于 FIFO 的任务队列调度方法，能够使短流程的并发任务共享堆栈存储空间，并且快速地进行切换，从而使 TinyOS 适用于并发任务频繁发生的传感器网络应用。当任务队列为空时，CPU 进入休眠状态，外围器件处于工作状态，任何外部中断都能唤醒 CPU，这样可以节省能量。而硬件事件处理线程，即中断处理线程，可以打断用户的轻量级线程和低优先级的中断处理线程，对硬件中断进行快速响应。

4）基于事件驱动模式的主动消息通信方式

每一个消息都维护一个应用层的处理程序。当目标节点收到这个消息后，就会把消息中的数据作为参数，并传递给应用层的处理程序，由其完成消息数据的解析、计算处理或发送响应消息等工作。

这种通信方式已经广泛应用于分布式并行计算。主动消息是并行计算机中的概念。在发送消息的同时，传送处理这个消息的相应处理函数和处理数据，接收方得到消息后可立即进行处理，从而减少通信量。由于传感器网络的规模可能非常大，导致通信的并行程度很高，传统的通信方式无法适应这样的环境。TinyOS 的系统组件可以快速地响应主动消息通信方式传来的驱动事件，有效提高 CPU 的使用率。

以上这些技术都是为了保证操作系统满足无线传感器网络的特殊要求，使其在处理能力和存储能力有限的情况下具有更强的网络处理和资源收集能力。

9.3.2　TinySEC 传感器网络安全体系结构

TinySEC 协议是在无线传感器网络上第一个完全实现链路层安全的体系结构，由 Berkeley 大学的 Chris Karlof、Naveen Sastry、David Wagner 等人研究设计，已经集成在 TinyOS 系统上，为传感器网络提供链路层的安全服务。

能够制定有效的安全机制是决定传感器网络广泛应用的主要因素之一。TinySEC 是针对传感器网络的轻量级安全机制，为传感器网络提供必需的安全服务组件。TinySEC 的优点在于能够很好地和传感器网络应用相融合，为各种应用提供安全支撑。

1. 安全体系结构的目标

传感器网络链路层安全协议需要达到的安全目标为：节点间的访问控制、数据完整性、数据保密性。TinySEC 不支持密钥管理，但是可以为应用层密钥管理模型提供安全传输服务。TinySEC 采用 RC5、Skipjack 等对称算法可以实现节点间的访问控制、数据完整性、数据保密性等安全服务，成为当前应用最广的传感网络安全协议。

1）访问控制

访问控制意味着链路层的协议应该能够阻止未授权的节点参与网络，合法的节点应该能够监测到未授权的节点并将其排除。

2）数据完整性

在数据的传输过程中，如果攻击者篡改了消息，接受者应该能够监测到这个改动。通过在每个数据包上添加信息认证码可以实现消息完整性认证。

3）数据保密性

保密性意味着保持数据对未授权节点的秘密性，数据的保密性通常是通过加密来完成的。为了保证传感器网络的安全特性，语义安全被实现，语义安全的实现意味着攻击者在对密文进行 yes 和 no 回答时，并不能够拥有高于 50% 的可能性。

2. TinySEC 安全体系结构的设计

现存的模式不适合传感器网络的安全需求。在现有的网络通信协议中，IPSec、SSL/TLS 和 SSH 都能满足 Internet 上的安全需求，在不安全的信道上建立安全的通信，保证通信的机密性。但是这些协议对于传感器网络来说开销太大。无线 Ad Hoc 网络的通信发展模式接近于无线传感器网络的需求，但是对于传感器网络来说，现存的设计模式是严格受限的。总结现存的模式，要么是不安全的，要么是不符合资源受限的传感器网络的，TinySEC 设计了一种适合于无线传感器网络的安全模式。

1）TinySEC 支持的安全选项

TinySEC 支持两种不同的安全选项。

- 认证加密（TinySEC－AE）模式。TinySEC 加密数据负载和使用 MAC 认证包。其中

MAC 由加密数据和信息包头计算得出。

- 仅有认证（TinySEC－Auth）模式。TinySEC 使用 MAC 认证整个包，但是数据负载是没有经过加密的。

（1）加密

使用语义安全加密需要两个设计决议：选择加密模式和指定初始向量 IV 格式。在 TinySEC 的设计中，使用指定的 8 字节 IV，使用密码分组链接（CBC）。下面将介绍TinySEC 中 IV 的格式及为什么 CBC 最适合于传感器网络的加密模式。

① TinySEC 的 IV 格式。

联系压缩安全带来的开销目标，IV 的长度和产生 IV 的方式可能对安全和性能带来很大的影响。如果 IV 太短，将会造成 IV 重复的危险。如果 IV 太长，将会增加数据包中不必要的位，造成很大的开销。根据鸽巢原理，无论如何选择 IV，n 比特的 IV 可以保证在发送 2^n+1 个数据包后重复。如果使用 n 比特的计数器在那个点之前重复将不会发生。但是对于一些 IV 产生策略，重复可能会产生早一些。要是选择 IV 位随机的 n 比特值，希望在发送 $2^{n/2}$ 个数据包之后出现第一次重复。

TinySEC 中 IV 的格式如图 9-9 所示。这里的 Dst 为接收者的目标地址，占两字节；AM 为操作类型的活动信息，占一字节；Len 为数据负载的长度，占一字节；Src 为发送者的源地址，占两字节；Ctr 为计数器值，占两字节。计数器从 0 开始计数，在每个数据包发送后计数值加 1。

Dst（2）	AM（1）	Len（1）	Src（2）	Ctr（2）

图 9-9　TinySEC 中 IV 的格式

② 加密模式。

首先说明为什么选择 CBC 模式作为传感器网络的加密模式。对称密钥有两种形式：流密码和使用分组密码的操作。流密码使用一个密钥 K 和 IV 作为一个种子，并衍生出一个伪随机密钥流 G_k(IV)。然后密钥流和信息进行异或操作：C =（IV，G_k（IV）$\oplus$P）。最快的流密码要快于最快的分组密码，这个性质使得流密码更加适合于资源受限的环境，但是流密码有一个失败的模型：如果两个不同的数据包使用同一个 IV 来进行加密，经常会导致明文的暴露。要保证 IV 不会被重复使用，就需要 IV 足够长，也就是说，至少 8 字节。既然传感器网络的目标之一就是最小化包的开销，则增加 8 字节到 30 字节的数据包是不被接受的。一个选择就是使用较短的 IV，接受 IV 的重复使用会发生。因此考虑一下原则：在目前的重复 IV 中，使用一个尽量具有顽健性的加密模式。很显然，流密码违背了这个原则，因此，仅有的选择就是使用分组密码的形式。

通常的分组密码是 8 位或 16 位，是一个应用在比特串上的伪随机变换。DES、AES、RC5 和 Skipjack 都属于分组密码。既然通常需要加密和认证长于 8 位或 16 位的信息，则分组密码需要一个加密更长信息的模式。对于 k 字节的分组密钥，一个典型的操作模式就是将信息拆分成长度为 k 字节的段，使用分组密码以一个指定的方式逐段加密每个分组。

由于大部分有效的消息认证码算法使用的是分组密码，节点需要在任何事件上实现分组密码，所以使用这个分组密码进行加密也节约了代码空间。这是使用分组密钥进行加密的优势。

如果使用分组密码来进行加密，必须选择一个操作模式，一个自然的选择就是选择 CTR 模式，但是 CTR 模型是一个流密码操作模型。另外一个自然的选择就是 CBC 模型。CBC 模型更好，主要体现在对于目前的重复 IV 退化得比较多。特别是在 CBC 模型下，如果使用相同的 IV 来加密两个明文信息 p 和 p′，密文将泄露最长 p 和 p′公共前缀的长度，而不会泄露更多。例如，如果 p 的第一个分组不同于 p′的第一个分组，通常密码分析者将不能够得到任何信息。因此，CBC 模式在重复 IV 的情况下仅泄露少量的信息，相对于流密码来说是一个很大的优点。在 IV 不重复的情况下，CBC 模式是可证明安全的。CBC 模型被设计使用一个随机的 IV，当使用一个计数器作为 IV 时，CBC 模型有一个分离的泄露问题。假设使用 IV 和 IV′分别加密两个明文 p 和 p′，如果 $p_1 IV = p_1' \oplus IV'$（这里的 p_1 表示 p 的第一个分组，其他的以此类推），在密文的第一个分组相等的情况下，将透露 p_1、p_1'。在一些情况下，可能泄露了部分明文。例如，假设 IV 是一个计数器值，并且 IV 和 IV′是两个连续的计数器值。经常有 $IV' = IV + 1$，如果明文偶尔满足了同样的模型，如 $p' = p + 1$，这样就会存在一个安全漏洞。这是不符合需求的。

幸运的是，简单确定允许 CBC 模型使用任何非重复的 IV。这个确定是预加密 IV，拒绝标准的 CBC 模型有利于这个变化。

使用 8 字节的分组密码来进行 CBC 模型加密多于 8 字节的密文。这个可能存在于信息的扩充中，从而增加了能量消耗。众所周知，使用技术确保密文的长度等同于明文的长度。加密少于 8 字节的有效载荷，产生的密文为 8 字节，因为密文泄露至少需要一个分组的密文。但是发送一个信息的开销（关闭通信设备、请求信道、发送开始符号）不利于缩短信息。

③ 分组密码选择。

传统的分组密码意味着当需要一个分组密码时，要么选择 AES，要么选择 3DES，但是 3DES 对于嵌入式的微处理器软件来说实现太慢，同样 AES 也是相当慢的。RC5 和 Skipjack 最适合于在嵌入式微处理器上进行软件实现。RC5 要稍微快一点，为了得到一个好的性能，RC5 需要一个被预计算的密钥模式，每个密钥需要使用 104 个额外字节的 RAM，由于这个原因，在 TinySEC 中，默认的分组密码为 Skipjack。

（2）消息完整性：需要 MAC

仅仅使用机密而没有认证是不安全的。没有认证机制来保证信息的完整性，接收者将不能够检测到信息在通信过程中是否被篡改。没有认证机制的信息也容易受到剪切－粘贴攻击。在剪切－粘贴攻击中，攻击者破坏部分未认证的加密信息，并构造出另外一些能够解密成具有一定意义明文的密文。例如，如果所有认证节点之间共享一个单密钥，攻击者能够从给一个节点的信息中析取被加密的数据负载，并且将这个负载发送给不同的节点，既然加密负载没有改变，第二个节点将成功解密并接收这条消息。

为了阻止这种攻击，TinySEC 总是认证信息，但是加密是可选的，仅仅当一些信息是敏感信息时，机密性才是必需的。对于不敏感的信息，加密不是必需的，而且还会增加响应时间、计算和能量开销。但是大部分的应用需要数据包认证，这意味着被认证的节点不会接收到攻击者注入的非法信息。

TinySEC 使用 CBC－MAC 结构来计算和认证 MAC。CBC－MAC 是高效和快速的，事实上其依赖于分组密码，CBC－MAC 是可证明安全的，但是标准的 CBC－MAC 对于可变尺寸

的消息并不是安全的。攻击者能够伪造特定消息的MAC。

2）数据包的格式

TinySEC数据包格式是基于目前TinyOS的。TinySEC数据包和TinyOS的不同如图9-10所示。共同的域是目的节点、活动信息类型（AM）和长度。活动信息类型类似于TCP/IP中的端口号。AM类型制定了适当的在接收端吸取和解释数据包的操作功能。这些域是不被加密的，因为用明文发送它们的好处超出了保护它们的秘密带来的好处。

Dst（2）	AM（1）	Len（1）	Grp（1）	Data（default length:29 bytes）	CRC（2）

（a）默认的TinyOS包结构

Dst（2）	AM（1）	Len（1）	Data（default length:29 bytes）	MAC（4）

（b）TinySEC-Auth包结构

Dst（2）	AM（1）	Len（1）	Src（2）	Ctr（2）	Data（default length:29 bytes）	MAC（4）

（c）TinySEC-AE包结构

图9-10　TinyOS默认和TinySEC协议的数据包格式对比

为了监测传输中的错误，TinyOS采用了CRC校验的方式。接收者计算校验码并和接收到的CRC进行比较，如果相等，则接收者接收数据包，否则丢弃。但是CRC不能监测到攻击者是否对数据包进行了恶意修改或伪造。为了保证信息的完整性和机密性，TinySEC用MAC来代替CRC，实现消息的完整性和可认证性。如图9-10（b）、（c）所示，TinySEC提供两种安全模式：TinySEC－Auth和TinySEC－AE。TinySEC－Auth仅通过MAC码实现数据包的鉴别，而TinySEC－AE具有数据加密和认证的功能，其包结构比TinySEC－Auth多4字节，主要用于构造CBC模式加密的初始向量IV。MAC可以保护整个数据包，包括目的地址、AM类型、长度、源地址、计数器及数据。这样防止了对数据包的篡改，同时也阻止了攻击者重定向数据包到另外一个节点，阻止了数据包丢弃和其他攻击。

参考文献

［1］李鸥，张效义，王晓梅，等. TinyOS实用编程：面向无线传感网节点软件开发. 北京：机械工业出版社，2013.

［2］李外云. CC2530与无线传感器网络操作系统TinyOS应用实践. 北京：北京航空航天大学出版社，2013.

［3］孙利民，李建中，陈渝，等. 无线传感器网络［M］. 北京：清华大学出版社，2006.

［4］Hill J，Szewczyk R，Woo A，et al. System architecture directions for networked sensors［C］//9th International Conference Architectural Support for Programming Languages and Operating Systems. Cambridge，MA，United states：Association for Computing Machinery，2000.

［5］Levis P，Madden S，Polastre J，et al. TinyOS：an operating system for wireless sensor networks［M］//Weber W，Rabaey J，Aarts E. Ambient Intelligence. New York，NY：Springer－Verlag，2005.

［6］Karlof C，Wagner D. Secure routing in wireless sensornetworks：attacks and countermeasures［C］//Sensor Network Protocols and Applications，2003：113－127.

［7］Bhatti S，Carlson J，Dai H，et al. MANTIS OS：an embedded multithreaded operating system for wireless

micro sensor platforms [C]. [S. l.]: Kluwer Academic Publishers, 2005: 563 – 579.
[8] 尹震宇，赵海，徐久强，等. 无线传感器网络操作系统中抢占式任务调度策略 [J]. 东北大学学报(自然科学版)，2007，28 (5)：652 – 655.
[9] ARM Limited. ARM security technology: Building a secure system using TrustZone technology. WhitePaper [EB/OL], http://www. realview. com. cn, Apr. 2009.
[10] 冯登国，秦宇，汪丹，等. 可信计算技术研究 [J]. 计算机研究与发展，2011.
[11] 陈书义，闻英友，赵宏. 基于可信计算的移动平台设计方案 [J]. 东北大学学报 (自然科学版)，2008. 9.
[12] 沈圣盛，基于 APB 总线的接口 IP 核设计与验证 [D]. 西安：西安电子科技大学，2011.
[13] 杜文银，张涛，凌君. 基于 TrustZone 技术的移动可信平台 [J]. 测控技术，2009.
[14] TIAGO ALVES, DON FELTON. TrustZone: Integrated Hardware and Software Security [EB/OL]. http://www. arm. com/products/esd/trustzone_home. html. 2004.
[15] ARM Limited. Cortex – A8 Technical Reference Manual Revision: r3p2 [M]. 2010.
[16] 周亦敏，隋伟鑫. ARM 架构中 TrustZone 安全处理技术的研究 [J]. 微计算机信息，2008.
[17] 魏兰. 基于 ARM TrustZone 的安全存储研究与实现 [D]. 成都：电子科技大学，2015.
[18] Tiago Alves, Don Felton. TrustZone：在嵌入式领域实现安全计算的软硬件集成保密系统 [EB/OL], http://www. realview. com. cn, 2007.

第 10 章　感知层无线接入网络安全技术

无线接入网络是连接物联网与互联网信息的枢纽点，对于信息的传递和网络的安全起着关键的作用，接入网安全技术将会从接入能力、管理能力、安全防护多个方面影响整个网络的性能。当前，接入网网络安全已经逐渐成为通信和计算机安全领域关注的热点，主流的接入技术有 WLAN、WiMAX、3G 及 LTE 四种，本章分别对这四类技术进行简要介绍概述，并针对每种网络的特点进行安全性分析，通过分析发现每种接入技术存在的安全隐患与问题，并在此基础上提出安全问题的解决方案。

10.1　无线局域网安全保密体系结构及实现

一个全方位的无线局域网安全保密体系结构包含无线局域网的物理平台安全、无线通信安全、无线网络运行安全等。需要充分利用各种先进的主机安全技术、身份认证技术、访问控制技术、密码技术、防火墙技术、安全审计技术、安全管理技术、系统漏洞检测技术、黑客跟踪技术等，在攻击者和受保护的无线局域网资源间建立多道严密的安全防线。

10.1.1　无线局域网安全目标

无线局域网安全保密是指保护无线局域网系统的硬件、软件和数据，防止在无线局域网上处理、存储或传输的数据的故意或偶然的泄露、更改、破坏或非授权访问，保障系统连续、可靠、安全地运行，其最终目标就是通过各种技术和管理手段实现无线局域网及其信息系统的保密性、完整性、认证性、可授权性、不可否认性等。另外，无线局域网的特殊性还要求安全方案具有可用性和高效性。

1. 保密性

保密性要求主要是指数据在无线局域网上处理、存储或传输的过程中，防止数据的未授权的公开和泄露，包括通信的隐蔽性、通信对象的不确定性和抗破译能力。其中通信的隐蔽性也叫对抗业务流分析。对于无线局域网系统而言，为了监听或窃取机密通信的信息，攻击者无须像在有线网络中实施攻击那样通过实际线路连接到通信信道上，而只需要在监控范围内架设一副天线就可获取到所需要的数据。

802.11 规范要求在物理层采用适当的传输措施，如采用直接序列扩频 DSSS、跳频扩频 FHSS、直接序列跳频扩频 FH/DS 等扩频技术。扩频技术具有很强的抗干扰性。要截获、窃听或侦察经过扩频技术处理的信号非常困难，除非采用与发送端所用的扩频码相同的伪随机序列且与之同步后进行相关检测，否则对扩频信号无能为力。数据加密也是无线局域网数据保密的一个重要环节。例如，在现有的无线局域网安全机制中，WEP 一般采用的是 RC4 流加密算法对网络中传输的数据分组进行加密，并采用有线等价保密，即协议来保护进入无线局域网中所需的身份验证过程。然而，直接序列扩频或跳频技术不能视为安全措施的主要部

分，因为这两种技术原本都只是通信技术，不能提供足够的抗恶意攻击的能力。而 WEP 自身存在严重的设计缺陷，存在密钥重用、弱密钥、IV 空间不足等安全问题，导致受其保护的无线局域网的保密性受到严重威胁。

2. 完整性

完整性是一种面向信息的安全性，它要求保持信息的原样，即要求在无线局域网中存储和传输的种类信息数据的精确性和可靠性，防止数据被恶意修改；对未授权数据修改操作进行自动检测和报警，并将所有操作行为记入日志。对存储和传输中的数据进行的修改包括改变、插入、删除、复制等。另一种潜在的修改是改变序列号和重放，这种修改一般在数据进入传输信道时发生。

现有无线局域网安全机制中采用 RC4 数据流加密算法来保护数据传输和提供数据完整性。已经证明，802. 11 规范对 RC4 加密算法及其密钥调度算法在实现中存在许多问题，使得攻击者可以预料到网络的秘密加密密钥，从而使得这些安全措施失效。另外，WEP 的完整性保护只应用于数据载荷，而没有包括应当保护的所有信息，如源、目的地址及防止重放等。对地址的篡改可形成重定向或伪造攻击，而没有重放保护可以使攻击者重放以前捕获的数据形成重放攻击。

3. 认证性

认证性是指实体提供了声称其身份的保证，以防止其他实体假冒。在通信过程中，认证性只对认证时的主体身份提供确认保证，为了获得认证的持续保证，需要将认证服务和数据完整性服务结合起来。

由于在同一网络中设置严密的用户口令及认证措施，可防止非法用户入网，因此当前 802. 11 网络认证性的中心问题是无线通信设备的身份认证，而不是用户身份验证或使用无线局域网的站点的身份验证。大多数无线 802. 11 产品都采用了密钥固定不变的共享密钥身份验证机制。802. 11b 定义了两种类型的认证服务：开放系统认证和共享密钥认证。开放系统认证是一种最简单的认证方案，用户只需要使用网络所分配的口令就可以进入网络。不过这种认证方案的安全系数不太高，但可以配合共享密钥认证来增强网络的安全性。

在提供可认证性服务时，需要防止针对该服务的以下两种类型的攻击：

① 对认证实体的重放攻击；

② 入侵者发起或响应的延迟攻击，从而造成的一种 DoS 攻击。

4. 可授权性

可授权性的目标是防止对无线局域网中的任何资源（计算资源、通信资源或信息资源等）进行非授权的访问。所谓非授权的访问包括未经授权的使用、泄露、修改、销毁及发布指令等。用户在获得这些授权之前必须进行身份识别，也就是认证。可授权性直接支持保密性、完整性、可用性及合法使用的安全目标，它对保密性、完整性和可用性所起的作用是十分明显的。

5. 不可否认性

不可否认性也叫不可抵赖性，是防止发送方或接受方抵赖所传输消息的一种安全服务。也就是说，当接收方接收到一条消息后，能够提供足够的证据向第三方证明这条消息的确来

自某个发送方，而使得发送方企图抵赖发送过这条消息的图谋失败。同理，当发送一条消息时，发送方也有足够的证据证明某个接收方的确已经收到这条消息。

6. 可用性

可用性是指可被授权的合法使用者在使用无线局域网时，系统应能提供完全满足使用者要求的各种信息和资源服务，而不会不合理地对这些要求进行拒绝或不能满足用户的合理要求。可用性攻击就是阻断信息和资源的合理使用，如破坏系统的合理运行就属于这种类型的攻击。

DoS 攻击以阻塞网络带宽、耗尽服务器内存资源、干扰及破坏正常通信为主，使网络的可用性受到严重威胁，在传统的有线网络中，DoS 已成为攻击者进行恶意破坏大型网站通信、破坏企业公众信誉形象、勒索讹诈公司资产的极具威胁性的途径。和有线网络一样，对于采用了高安全级别设定的无线网络环境，攻击者也可以使用多种方式对无线接入点进行拒绝服务攻击，这些攻击多数都可导致无线接入点服务中断、过载，无线网络丢包率增大，甚至出现无线接入点假死需要重启的状况。

7. 高效性

由于无线局域网各种资源的限制，所设计的安全方案如果需要在可接受的时限内提供所期望的安全服务，就需要考虑高效原则。例如，协议交互的消息数目应尽可能少，终端完成的任务应尽可能少，尽量采用专用安全防护器件实施保护，充分利用先前已经建立的信任关系，以减少再次认证成本等。

10.1.2　主要安全威胁

无线局域网的综合防御体系需要从平台安全、通信安全、运行安全等多个层面进行构建。下面分析各个层面所面临的安全威胁。

1. 物理平台安全威胁

无线局域网中的物理平台设备主要包括站点（STA）和接入点（AP）两类。如前所述，站点通常由一台 PC 或笔记本电脑加上一块无线网络接口卡构成；接入点通常由一个无线输出口和一个有线的网络接口构成，其作用是提供无线和有线网络之间的桥接。物理平台安全威胁是关于这些无线设备自身的安全问题，主要表现在以下几方面。

① 无线设备存在许多限制，这将对存储在这些设备中的数据和设备间建立的通信链路安全产生潜在的影响。与个人计算机相比，无线设备如个人数字助理等存在电池寿命短、显示器小等缺陷。

② 无线设备虽有一定的保护措施，但这些保护措施总是基于最小信息保护需求的。如果存储重要信息的无线设备被攻击者成功入侵，则攻击者就可能无限期地对设备拥有唯一的访问权，不断地获取受保护的数据。

因此，必须加强无线物理平台设备的各种防护措施。

物理平台的构成主要是组成系统各部分的软/硬件支撑系统，它们构成了无线局域网的基本计算环境。无线局域网中的物理平台设备面临的威胁主要有以下几方面。

① 物理硬件平台缺乏完整性保护和验证机制，平台中各个模块的硬件容易被攻击者篡改；平台内部的各个通信接口缺乏机密性和完整性保护，在此传递的信息容易被窃听或篡改；平台

缺乏完善的访问控制机制，特别是基于硬件的强访问控制，信息容易被非法访问和窃取。

② 各个不同的物理平台可能会使用不同的操作系统软件，这些操作系统可能并不安全，存在许多公开的漏洞。

③ 物理平台所支持的各类无线应用自身可能存在固有的安全隐患和程序漏洞，这给计算能力有限的无线物理平台带来了更大的安全威胁。同时，这些无线应用自身也可能会丰富无线物理平台感染病毒、木马和蠕虫的渠道。

④ 用户对这些物理平台的配置能力逐渐增强，不合理的配置很可能会导致安全级别的降低。随着性能的提升，物理平台自身可能会被利用成为新的攻击工具，用于入侵无线网络和与之相连的有线网络。

⑤ 传统防病毒、入侵检测等软件的体积将随着病毒种类的增加和入侵手段的日新月异而不断增大，并不适合计算、存储能力及电池容量有限的无线物理平台。

2. 无线通信安全威胁

无线局域网的传输介质的特殊性使得信息在传输过程中具有更多的不确定性，受到的影响更大，主要表现在以下几方面。

① 窃听。任何人都可以用一台带无线网卡的 PC 或廉价的无线扫描器进行窃听，但是发送者和预期的接收者无法知道传输是否被窃听，且无法检测窃听。

② 修改替换。在无线局域网中，较强节点可以屏蔽较弱节点，用自已的数据取代，甚至会代替其他节点做出反应。

③ 传递信任。当机构的网络环境中包括开放（安全强度弱）无线接入点时，就会为攻击者提供一个不需要物理安装的开放接口用于网络入侵。因此，要求防止网络环境中存在安全性弱的接入点，或对这类接入点进行屏蔽。

虽然组成无线局域网的软/硬件平台对外提供的交互接口较少，相对专用且不同网络内部的接口类型单一，但是无线局域网传输媒介的开放性使得系统同样面临来自网络的攻击威胁，如漏洞攻击、特洛伊木马攻击、病毒攻击等。网络上传输的信息也面临着信息泄露的安全威胁。外部系统通过网络对系统内部服务与数据的访问存在身份假冒和越权访问的安全威胁。对无线通信链路的安全威胁还包括信息流量分析、对信息流的截断/篡改/插入/重放等攻击，以及各种针对通信协议缺陷的入侵攻击和拒绝服务攻击等。例如，攻击者通过对信息流向、流量、通信频度和长度等参数的分析，猜测对方的意图，获取有用信息，就可对系统造成极大的数据安全威胁。

3. 无线网络运行安全威胁

无线局域网作为一个软/硬件集成的有机的网络整体，其安全性除包括这些软/硬件设备自身的安全保护能力，以及无线网络通信安全以外，还包括围绕无线局域网系统的运行过程和运行状态的安全，主要涉及信息系统的正常运行与有效的访问控制等方面的问题，面对的威胁包括网络攻击、网络病毒、网络阻塞、系统安全漏洞利用等。进行保护，建立集保护、检测、响应和恢复于一体的全面、灵活的动态运行安全保障体系也十分重要。

无线局域网运行安全面临的威胁有以下几方面。

① 基础结构攻击。基础结构攻击基于系统中存在的漏洞，如软件 bug、错误配置、硬件

故障等。针对这种攻击进行保护几乎是不可能的，所能做的就是尽可能地降低破坏所造成的损失。

② 拒绝服务。无线局域网存在一类比较特殊的拒绝服务攻击，攻击者可以发送与无线局域网相同频率的干扰信号来干扰网络的正常运行，从而导致正常的用户无法使用网络。

③ 置信攻击。通常情况下，攻击者可以将自己伪造成基站。当攻击者拥有一个很强的发送设备时，就可以让移动设备尝试登录到他的网络，通过分析窃取密钥和口令，可发动针对性的攻击。

10.1.3　无线局域网安全需求

对于一个无线局域网系统来讲，其安全目标包括保密性、完整性、认证性、可授权性、不可否认性、可用性和高效性等，根据其安全目标和各个层面所存在的安全威胁，可以提出如下分层安全需求。

1. 平台安全需求

无线局域网中的物理平台安全需求主要包括平台计算环境的健壮性和可信性需求、关键部件的完整性和正确性需求、恶意代码防范体系需求、身份认证需求、访问控制需求、数据的机密性和完整性需求等多个方面。其中，平台计算环境的健壮性和可信性需求是平台安全的基本需求之一，是有效降低系统节点所面临的软/硬件风险的前提。身份认证需求要求必须对系统用户或软/硬件实体的身份进行安全验证，它也是其他安全措施的基础。数据的机密性和完整性需求是保护平台存储数据安全，确保整个系统的信息安全的重要保证。

平台安全具体的安全需求包括以下几方面。

① 对无线网络用户和管理员进行高可靠的身份认证。

② 控制系统操作人员的安全等级，明确划分不同的管理权限，实行强制访问控制。

③ 采用可信移动平台的思想，添加可信启动、完整性检验和保护存储等措施，保证物理平台硬件、操作系统和应用软件的完整性、可靠性，以及其配置信息的合法性，防止计算机病毒木马、后门和其他恶意代码的攻击。

④ 保护存储在平台中的敏感数据的机密性、完整性，并能够及时恢复受损的关键数据。

⑤ 在进行敏感信息交互之前，确保移动平台当前状态的可信性，并保护发送数据及数据流的机密性和完整性，验证接收数据的完整性和真实性。

⑥ 增加物理平台的集成度，减少可被攻击的硬件接口，确保物理平台中各个接口上传递数据的完整性和机密性。

⑦ 对向外发送的数据进行内容扫描，检测可能出现的违反安全保密策略的事件或内容。

⑧ 对出入系统的连接，包括服务连接请求和对外访问请求等进行检测与控制，防止非授权的连接。

⑨ 应防止物理平台的丢失而带来的危害，阻止物理平台被非法滥用或作为发起攻击的工具。

⑩ 增强物理平台驱动程序抵抗攻击的能力。

2. 无线通信安全需求

无线局域网系统的通信安全需求主要包括通信链路的健壮性需求、通信数据的机密性和

完整性需求等。

通信链路的健壮性需求是指网络要能够在恶劣的环境和敌人干扰的情况下保证通信链路的畅通。无线局域网的通信特点决定了通信链路容易受到来自自然环境和人为的干扰，必须保证在受到（敌方）干扰的情况下，通信链路能正常、持续地运行。很多相关技术可实现这一要求，如扩频技术、跳频技术和智能天线技术等。

通信数据的机密性和完整性需求具体包括以下几方面。

① 在用户、接入网络和归属网络之间实施双向认证。由于无线设备资源受限，因此应尽可能采用开销较小，但却安全可靠的认证机制。

② 确保信令数据和用户业务数据在无线及有线接口的机密性。

③ 确保信令数据、用户业务数据及控制数据的完整性和数据源认证。

④ 对出入的连接请求和服务访问请求实施控制，禁止非法的访问和连接。

⑤ 对传输的数据进行监测，包括对病毒、木马等恶意代码的检测，以及对违反安全保密策略的事件或内容的报警和过滤。

⑥ 提供安全有效的密钥协商机制，以及可靠的密钥备份和恢复机制。由于无线设备资源受限，因此应尽可能采用开销较小，但却安全可靠的密钥协商机制。

⑦ 确保漫游切换过程中，用户、接入网络和归属网络之间的认证性、机密性、完整性不因切换而受影响。

3. 无线安全运行需求

由于无线局域网系统的结构复杂，单纯依靠一种或几种安全技术难以有效地解决系统的安全问题，必须围绕信息对抗原理建立一个完整的无线网络安全防护体系。该体系应具有良好的普适性，能够适应各种复杂的情况；具有可扩展性，能够较容易地集成新的安全技术以提升系统安全。这些特点要求系统在安全体系结构设计和运行安全方面要特别考虑。

安全运行防护体系必须采用分布式的纵深防御结构，建立集保护、检测、响应和恢复于一体的全面、灵活的动态运行安全保障体系，突出系统各部分安全防护的自主性和多层多级保护特性；系统各组成部分之间必须要有设计合理、功能明确的安全协作接口；所构建的安全体系必须具有开放性和灵活性，能够在统一的框架下分析和解决系统出现的新问题，融合未来出现的新技术。

具体的安全需求包括以下几方面。

① 监控无线局域网的网络边界及其内部的无线信号，对信号进行解析。

② 对来自外部的恶意入侵行为，包括扫描行为、测试行为和渗透行为进行监测和控制，必要时采取诱骗、跟踪或反制措施。

③ 能够及时检测出对系统的非法扫描、探测行为并予以报警。

④ 能够及时检测出对无线设备平台、信息链路或信息服务的拒绝服务攻击，并能采取措施定位、跟踪、隔离攻击源。

⑤ 能够及时检测出对信息节点的渗透攻击行为，并能及时报警和采取隔离、恢复措施。

⑥ 能够有效评估攻击威胁的大小、攻击的目的及其可能造成的危害。

⑦ 能够将检测到的信息及时反馈给安全控制系统或控制设备，实施有效的控制。

⑧ 能够对网络拓扑实施有效的检测和控制。

⑨ 能够检测出局域网内部存在的威胁节点，并能对其进行定位、隔离和报警。

10.1.4　需要的安全措施

上述无线局域网安全需求需要通过各种具体的安全措施，由种类具体的安全保障系统、安全产品和安全组件相互配合来加以实现。这些安全措施是实现无线局域网安全体系结构所必需的，并由此构成如图 10-1 所示的系统安全功能体系和安全结构体系。综合这些安全措施，可形成无线局域网安全保密与防护的整体解决方案，能够提供对无线局域网从无线终端到无线网络，再到网络运行的强有力的安全保密和防护保证。

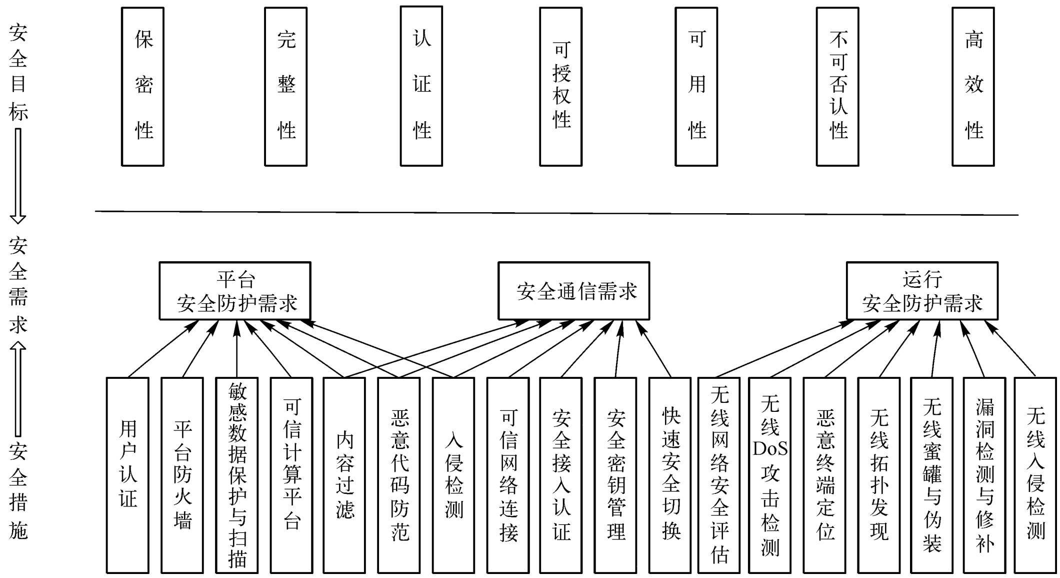


图 10-1　系统安全功能体系和安全结构体系

10.1.5　安全无线局域网的基本结构和实现方案

1. 网络基本部署结构

安全无线局域网可在很多类型的应用场景中得到应用。为了便于分析，这里给出一个通用的应用场景，体现了安全无线局域网的大多数结构特征，主要涉及无线终端、无线接入点、接入控制器、AAA 服务器/AAA 代理等关键部件，其构成如图 10-2 所示。

1）无线终端（STA）

无线终端主要指的是用于接入安全无线局域网的各类信息终端，如笔记本电脑、手机或掌上电脑等。其主要功能如下。

- 无线网络搜寻功能：可以根据用户的要求来选择接入的无线局域网。这里的无线局域网一般由 SSID 标识。
- 无线网络连接功能：无线终端提供的客户端软件允许用户建立一个配置文件以存放他们的无线网络选择并自动建立连接。

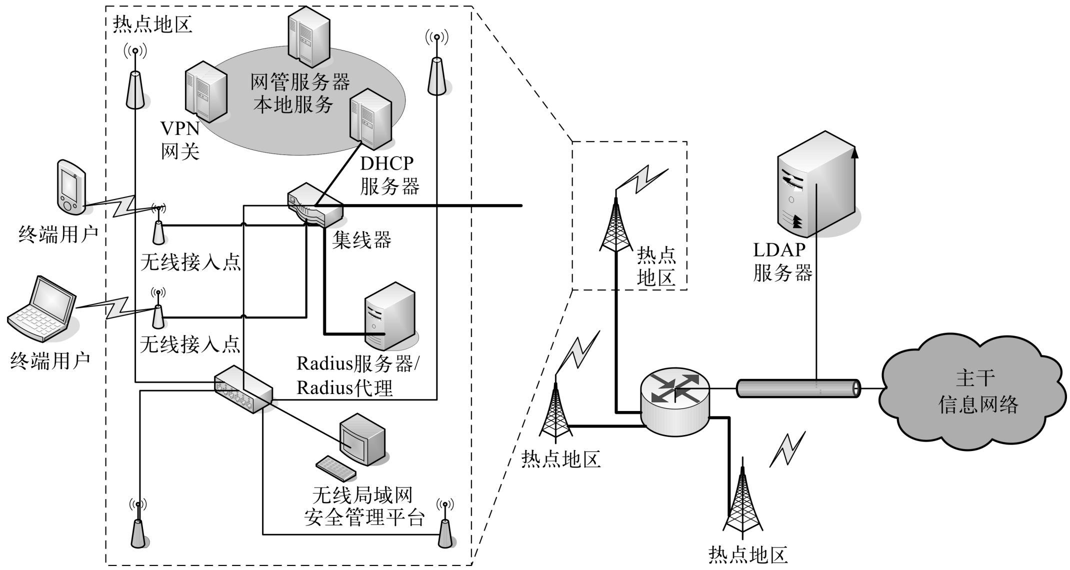


图 10-2　安全无线局域网的基本部署结构

- 无线网络认证功能：无线终端通过 802.1X 协议或其他的认证方式，如基于浏览器的认证方式进行用户认证。
- 无线网络连接管理：无线终端根据用户的设置维护无线网络的连接。

2）无线局域网安全接入点（AP）

AP 是无线移动终端和固定网之间数据交换的桥梁，同时也是无线移动终端接入无线局域网的受控端。AP 具有如下功能。

- 提供基于端口的访问控制功能，配合认证服务器端对用户的接入进行控制，阻止非法客户端的接入。
- 认证功能：接入过程中的安全协议及密钥分配等由一系列状态机实现。按照协议的设计流程，状态机基于当前的状态和收到的信息进行运算，根据结果判断转入下一个状态。
- RADIUS 客户端功能：本系统采用 RADIUS 服务器作为认证授权服务器。AP 是 RADIUS 服务器的客户端，控制无线接入的端口。AP 上的 RADIUS 客户端实现与 RADIUS 服务器的协议交互。
- 提供无线接入网与固定网之间的网桥功能，把将要访问控制的局域网与主干信息网络连接起来。
- 基于链路层的数据加密功能。
- 提供移动用户漫游切换功能，确保用户在不同 AP 覆盖范围之间切换时，能提供较好的通信服务质量。

除了以上一些基本功能外，AP 还具有以下功能。

- 认证信息转发功能：如提供基于 802.1X 的接入控制功能，把无线终端的认证信息转发给 AAA 服务器。
- 提供 AAA 服务器的功能：即 AP 不需要连接认证服务器，独自对用户进行认证。

- 用户 MAC 地址接入控制、用户二层隔离等功能。
- 可提供 VLAN、DHCP 等服务。

具有基本功能的 AP 称为“瘦”AP，具有以上功能的 AP 称为“胖”AP。

3）接入控制器服务

接入控制服务位于无线接入网与主干信息网络的交界处，主要提供无线终端连接主干信息网络的接入控制服务。它的主要功能包括：

- 充当主干信息网络的网关；
- 提供如 DHCP、DNS、VLAN、VPN 等本地网络服务；
- 实施对网络的业务控制；
- 用户管理控制；
- 网络数据的监控；
- 网络设备的管理。

4）AAA 服务器/AAA 代理

AAA 服务器/AAA 代理主要向用户提供 AAA 功能，或者向那些漫游用户提供 AAA 代理功能。AAA 服务器通过 EAP 和无线终端进行认证，并生成 WPA 会话密钥分发给 AP 和无线终端。

5）无线局域网安全管理

对无线局域网空间进行实时监控，针对无线局域网中出现的各种安全威胁，进行预防和防御处理，以提高无线局域网的安全性能。其主要功能如下。

- 无线局域网拓扑发现功能：对无线局域网的无线网络流量进行实时捕获，通过对捕获报文进行分析，快速分析和反映出无线网络拓扑结构，以实现对无线网络环境的实时监测。
- DoS 攻击探测功能：目前无线局域网安全措施还不能很好地应对 Authentication Flood、Deauthentication Flood、Beacon Flood 攻击，DoS 攻击探测对这几种攻击进行探测，如果发现存在 DoS 攻击，则采取相应的防护措施。
- 无线入侵检测功能：根据无线局域网所受安全威胁的特征，结合传统入侵检测技术，对无线局域网中的入侵行为进行有效的防范。
- 无线定位功能：无线局域网入侵者采用无线入侵的方式，检测到攻击后需要对攻击者进行物理定位，可及时解除存在的安全威胁。无线定位由定位处理和定位信息采集点组成，定位信息采集点采集无线局域网信号，定位处理单元根据采集的数据结合信号特征数据对特定特征的无线设备进行定位。
- WEP 和 WPA－PSK 密钥强度攻击、DoS 攻击及安全评估功能：无线局域网目前存在的攻击有 WEP 密钥攻击、WPA－PSK 密钥攻击等，对无线局域网进行这两种攻击检测，最终评估出无线局域网的防御能力（密钥强度）。另外，对于对无线局域网的 Beacon Flood、Deauthentication Flood、Authentication Flood 攻击，根据攻击对网络的影响评估无线局域网的抗 DoS 攻击的能力。
- 系统漏洞扫描：接入无线局域网中的合法实体也会给整个局域网带来安全威胁。合法移动终端可能存在安全漏洞，这些漏洞很容易被恶意软件攻击，并对整个无线局域网系统造成安全威胁。系统扫描对局域网中所有的实体进行漏洞扫描，并对存在漏洞的

实体进行相应的安全处理。

- 无线蜜罐功能：入侵者渗透到局域网中，其中一个目的是要访问局域网的网络资源。无线蜜罐可以迷惑已经接入无线局域网中的攻击者，从而为解除入侵威胁赢取缓冲时间。
- 威胁处理功能：威胁处理对无线局域网的接入设备进行管理，若检测到攻击，平台就进行报警，对入侵者进行定位，并采用关闭接入点的方式解除入侵者的连接；另外，对无线网络通信记录进行安全审计，以便分析和发现未知的异常。

其中，无线终端（STA）在配备安全无线网卡之后，与无线安全接入点、RADIUS 服务器和 LDAP 服务器一起，构成基本的无线局域网安全接入部分，另外加上无线局域网安全管理平台，构成安全无线局域网。其基本的硬件部分构成如图 10-3 所示。

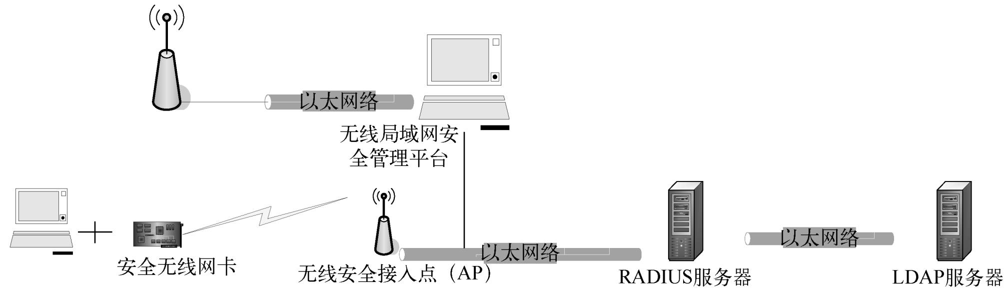


图 10-3　安全无线局域网硬件部分构成图

2. 网络基本实现方案

上述这些部件中，涉及硬件设备设计开发的主要有无线终端（STA）的无线网卡和安全接入点（AP），RADIUS 服务器、LDAP 服务器、无线局域网安全管理平台是在主机中采用软件实现的，不涉及硬件部分。方案中的 STA、AP、RADIUS 服务器、LDAP 服务器、无线局域网安全管理平台部分实现的结构框图如图 10-4 所示。

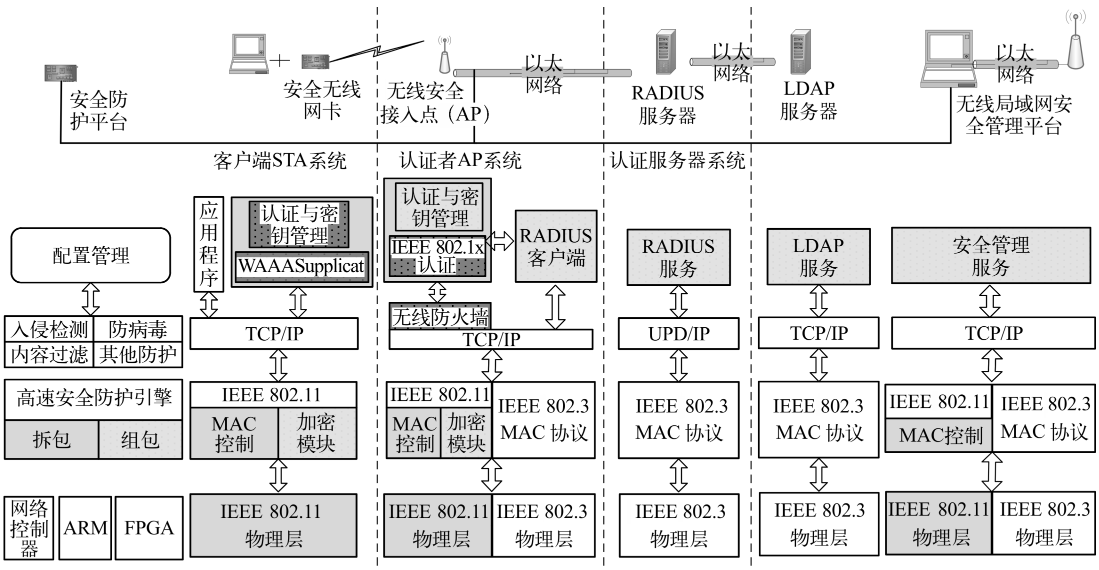


图 10-4　安全无线局域网各部分实现方案的结构框图

在图 10-4 中，阴影部分为安全无线局域网中要实现的模块，其中的颜色较深部分为安全技术实现模块。802.11 PHY 和 802.3 PHY/MAC 采用现成的网络控制和收发芯片来实现，802.11 MAC 层和加密模块在嵌入式微处理器内部实现，而认证协议和密钥管理模块采用软件实现。安全防护平台基于 ARM + FPGA 的架构，能对网络数据包进行拆包组包操作，实现高速多模式匹配，满足各类安全应用需求。

10.2　WiMAX 安全技术

10.2.1　WiMAX 网络概述

1. 标准规范

WiMAX 是 IEEE 802.16 标准在市场推广方面采用的名称，同时因为它也是一项基于 IEEE 802.16 系列标准的宽带无线接入城域网（BWAMAN）技术，所以常被称为 IEEE 无线城域网（WMAN）。IEEE 802.16 工作组是 WiMAX 空中接口标准的制定者，主要针对 WMAN 的物理层和 MAC 层制定规范和标准。IEEE 802.16 技术的最初目的是提供最后 1km 宽带无线接入技术。为了服务更广阔的市场，IEEE 802.16 标准转变该技术的焦点到一个更加类似蜂窝无线通信系统的移动架构上。如今 WiMAX 已经成为一个能够适应市场需求，提供增强用户移动性的多功能技术。

IEEE 802.16 工作组先后发布了 IEEE 的 802.16—2001、802.16a、802.16c、802.16d、802.16e、802.16f、802.16g、802.16h、802.16i、802.16j、802.16k、802.16m、802.16n 和 802.16p 等系列标准。其中主要标准的演进路线如图 10-5 所示。

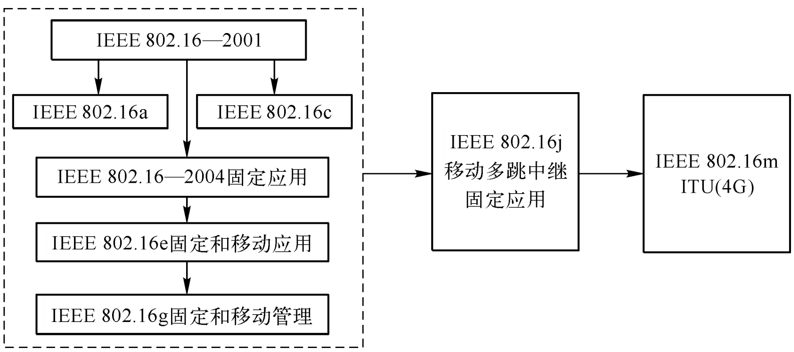


图 10-5　IEEE 802.16 主要标准的演进路线

IEEE 802.16—2001 标准于 2001 年 12 月发布，对 10 ~ 66GHz 固定宽带无线接入系统的空中接口物理层和 MAC 层进行了规范，由于使用的频段较高，因此仅能应用于视距范围通信。IEEE 802.16a、IEEE 802.16c 两个标准分别于 2002 年和 2003 年发布，主要对 IEEE 802.16—2001 标准的缺陷进行了修正，改善了互操作、服务质量等方面的性能。

IEEE 802.16d 标准于 2004 年发布，该标准冻结版本称为 IEEE 802.16—2004，它详细规范了 2 ~ 66GHz 固定宽带无线接入系统的空中接口物理层和 MAC 层，整合前期规范，引入非视距传输，引入 OFDM 技术，在 20MHz 的信道范围内提供 75 Mbps 的速率。IEEE

802.16d 标准也被称为固定 WiMAX。它是 IEEE 802.16 系列标准中相对比较成熟且第一个具有实用性的标准版本。

作为 IEEE 80.16—2004 标准的修正案，IEEE 802.16e 对固定无线网络业务进行了增强，并且具备了蜂窝移动通信能力。IEEE 802.16e 标准又称移动 WiMAX，于 2005 年 12 月通过，工作频率小于 6GHz，支持固定、机动、便携环境，并最终支持 120km/h 的移动环境。为了支持移动性，IEEE 802.16e 在 IEEE 802.16d 的基础上增加了切换支持、节电的睡眠模式、寻呼及增强的安全能力等。

2009 年 5 月，IEEE 将 802.16d—2004、802.16e—2005、802.16—2004/Corl—2005、802.16f—2005 和 802.16g—2007 等标准合并成最新的 IEEE 802.16—2009 标准。此后，IEEE 于 2009 年 6 月正式发布了 IEEE 802.16j—2009，该标准在 IEEE 802.16—2009 的基础上进行了扩充，主要是增加了中继（Relay）和多跳（Multihop）功能，从而提高了网络吞吐量和增大了网络覆盖范围。为了应对来自其他标准组织的挑战，顺利成为 IMT－Advanced（4G）标准之一，IEEE 标准委员会于 2006 年 12 月通过了 IEEE 802.16 工作组提交的 IEEE 802.16m 标准的立项申请。IEEE 802.16m 项目的主要目标有两个：一是满足 IMT－Advanced 的技术要求，二是保证与 IEEE 802.16e 兼容。2009 年 10 月，国际电信联盟正式确定 IEEE 802.16m 与 LTE－Advanced 一起作为 IMT－Advanced 国际标准候选技术，2012 年 1 月，WirelessMAN－Advanced（802.16m）和 LTE－Advanced 技术规范被正式确立为 4G 国际标准。JEEE 802.16 系列标准规范见表 10-1。

表 10-1　IEEE 802.16 系列标准规范

标　准　号	发布时间	主要内容
IEEE 802.16—2001	2001	10～66GHz 固定宽带无线接入系统空中接口标准
IEEE 802.16a	2003	2～11GHz 固定宽带无线接入系统空中接口标准
IEEE 802.16c	2003	10～66GHz 固定宽带无线接入系统关于兼容性的增补文件
IEEE 802.16d	2004	固定宽带无线接入系统空中接口标准（10～66GHz、＜11GHz）
IEEE 802.16e	2005	固定和移动宽带无线接入系统空中接口标准（＜6GHz）
IEEE 802.16f	2005	固定宽带无线接入系统空中接口管理信息库（MIB）要求
IEEE 802.16g	2007	固定和移动宽带无线接入系统空中接口管理平面流程和服务要求
IEEE 802.16h	2010	将免许可频段的研究从固定系统推进到移动系统
IEEE 802.16i	2008	定义 IEEE 802.16 移动接入系统 MAC 层和物理层的 MIB 及相关管理流程
IEEE 802.16j	2009	增加中继和多跳功能，提高网络吞吐量和增大网络覆盖范围
IEEE 802.16k	2007	针对 IEEE 802.16 桥接进行修改，使之能与 IEEE 802.16MAC 兼容
IEEE 802.16m	2011	满足 IMT－Advanced 技术要求，保证与 IEEE 802.16e 兼容
IEEE 802.16n	2013	使 IEEE 802.16 标准适应高可靠网络
IEEE 802.16p	2012	使 IEEE 802.16 标准适应机对机（Machine－to－Machine）应用

2. 网络拓扑

IEEE 802.16 支持两种网络拓扑结构，即 PMP（点对多点）模式和 Mesh 模式。

1）PMP 模式

点对多点网络拓扑结构由一个核心基站（BS）和多个用户站（SS）或者移动台（MS）组成，通常被用于最后 1km 宽带接入、私人企业到远距离办公室的连通，以及多站点的长距离无线通信服务。PMP 网络能使用视距（LOS）或非视距（NLOS）信号传输。其具体网络拓扑结构如图 10-6 所示。

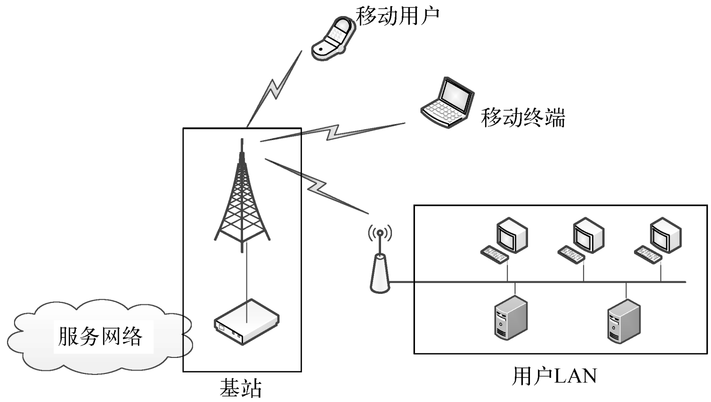


图 10-6　点对多点网络拓扑结构示意图

2）Mesh 模式

Mesh 网络拓扑是 PMP 结构与自组织网络的结合，是一种高容量、高速率的分布式无线网络结构。与 PMP 网络拓扑不同的是，Mesh 网络拓扑中的每个 SS/MS 都具备路由选择的功能，SS/MS 既是业务的使用者又是业务的提供者，即它具有数据的转发功能，可以向网络中的其他节点转发它所接收到的数据包。多个 BS 之间可以通过中继站（RS）方式来实施数据转发，从而扩展整个网络的地理覆盖范围。Mesh 拓扑结构如图 10-7 所示。

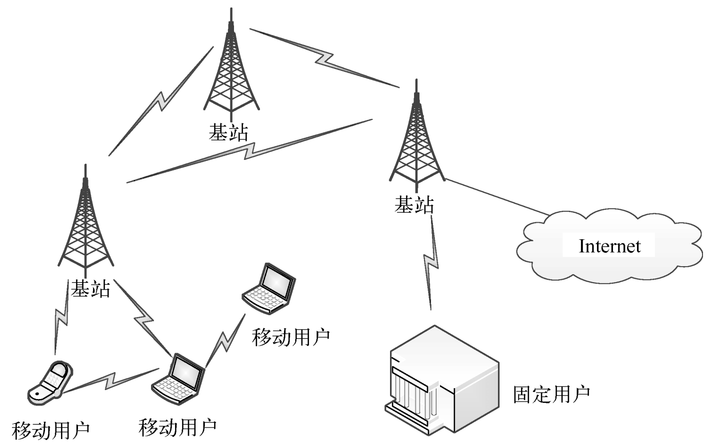


图 10-7　Mesh 拓扑结构示意图

3. 协议架构

IEEE 802.16 标准协议框架（见图 10-8）定义了介质访问控制（MAC）和物理层（PHY）协议结构。PHY 层支持在一个较宽频谱范围（2～66GHz）内的灵活操作，包括信

道带宽，频分复用、时分复用等。MAC 层定义了不同物理层上的通用特性，功能性覆盖了初始搜索、网络接入、带宽请求、面向连接管理，以及整个动态 WiMAX 环境的信息安全。

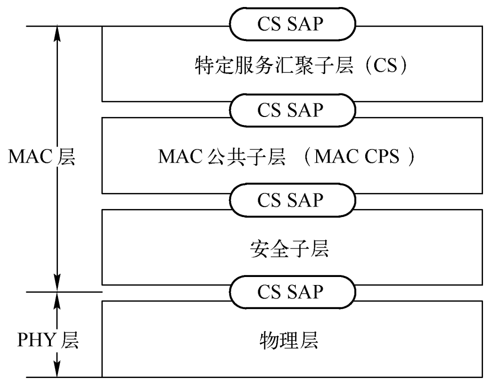


图 10-8　IEEE 802.16 标准协议框架

1）MAC 层

WiMAX 中的通信是面向连接的。来自 WiMAX MAC 上层协议的所有服务（包括无连接服务）被映射到 WiMAX MAC 层 SS 与 BS 间的连接。为向用户提供多种服务，SS 与 BS 之间可以建立多个连接，并通过 16bit 连接标识（CIDs）识别。

MAC 层又分为特定服务汇聚子层（CS）、MAC 公共子层（CPS）和安全子层（SS）三个子层。

① 特定服务汇聚子层（CS）。CS 子层主要负责与高层接口，将所有从汇聚层服务接入点（CS SAP）接收到的数据转化/映射成 MAC 服务数据单元（SDU），并通过 MAC 接入点（MAC SAP）发送给 MAC 公共子层，关联到正确的 MAC 服务流（SFID）及连接（CID）。协议提供了多个 CS 规范作为与外部各种协议的接口，可实现对异步传输模式（ATM）、IP 等协议数据的透明传输。

② MAC 公共子层（MAC CPS）。CPS 子层实现 MAC 层的核心功能，包括系统接入、带宽分配、连接建立和连接维护等。它通过 MAC 层 SAP 接收来自各种 CS 层的数据，并分类到特定的 MAC 连接，同时对物理层上传输和调度的数据实施 QoS 控制。

③ 安全子层。安全子层主要实现认证、密钥交换和加解密处理等功能，直接与 PHY 交换 MAC 协议数据单元（MPDU）。安全子层内容较多，包括密钥管理（PKM）协议、动态安全关联（SA）产生和映射、密钥的使用、加密算法、数字证书等。该子层支持 128bit、192bit 及 256bit 加密系统，并采用数字证书的认证方式，以保证信息的安全传输。

2）PHY 层

PHY 层由传输汇聚子层（TCL）和物理媒体相关（PMD）子层组成，通常说的物理层主要是指 PMD。IEEE 802.16 物理层定义单载波（SC）、SCa、OFDM、OFDMA 四种承载体制，以及 TDD 和 FDD 两种双工方式。上行信道采用 TDMA 和 DAMA 体制，单个信道被分成多个时隙，SS 竞争申请信道资源，由 BS 的 MAC 层来控制用户时隙分配；下行信道采用 TDMA 体制，多个用户数据被复用到一个信道上，用户通过 CID 来识别和接收自己的数据。

10.2.2　WiMAX 安全体系架构

WiMAX 安全有两个目标：一是提供无线网络数据传输机密性保护；二是提供网络接入控制。机密性保护通过对 MS 和基站之间的无线链路进行加密来实现，基站通过仅提供加密业务服务来防止非授权用户接入。WiMAX 同时制定了密钥管理协议来提供密钥分发和身份认证服务，用于实现 BS 与 MS 之间的密钥同步及基于证书的访问控制。

IEEE 802.16 安全子层的协议架构如图 10-9 所示，主要由加密封装协议和密钥管理协议两类协议组成。加密封装协议主要为各类协议数据单元提供加解密服务，而密钥管理协议则主要为 SS 提供密钥分发服务。

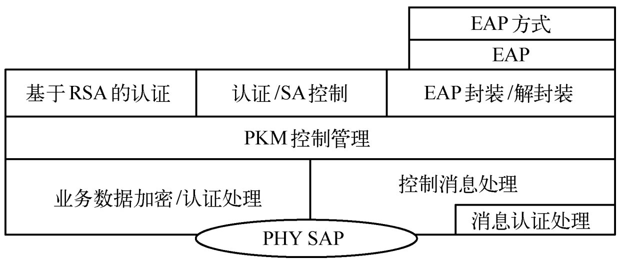


图 10-9　安全子层的协议架构

安全子层协议各模块功能如下。

- PKM 控制管理：控制所有安全组件，各种密钥在此层生成。
- 业务数据加密/认证处理：对业务数据进行加解密，执行业务数据认证功能。
- 控制消息处理：处理各种 PKM 相关 MAC 消息。
- 消息认证处理：执行消息认证功能，支持 HMAC、CMAC 或 short - HMAC。
- 基于 RSA 的认证：当 SS 和 BS 之间的认证策略选择 RSA 认证时，利用 SS 和 BS 的 X.509 数字证书执行认证功能。
- EAP 加密封装/解封装：提供与 EAP 层的接口，在 SS 和 BS 认证策略选择基于 EAP 的认证时使用。
- 认证/SA 控制：控制认证状态机和业务加密密钥状态机。

1. 数据加密协议

数据加密协议主要为宽带无线网络上传输的分组数据提供机密性、完整性等保护。

数据加密协议定义了加解密算法、认证算法及密码算法应用规则等一系列密码套件。IEEE 802.16—2004 仅支持 DES - CBC 加密算法（此算法已是不安全的），IEEE 802.16e 和 IEEE 802.16—2009 同时支持 DES - CBC 及 AES - CBC、AES - CTR、AES - CCM 等 AES 数据加密模式，而 IEEE 802.16m 标准仅支持 AES 数据加密模式。

WiMAX 安全子层仅对 MAC PDU 的载荷部分数据进行加密，对 MAC PDU 的头部、子头及携带的管理消息均不加密。这种处理方式虽然方便了各种 MAC 层操作，但是也带来了安全隐患，致使 MAC 层的各种管理消息很容易被截获分析。

2. 密钥管理协议

密钥管理协议用于实现由 BS 向 SS 安全分发密钥数据。BS 和 SS 通过 PKM 实现密钥同步，而且 BS 还可利用 PKM 实现网络接入控制。目前有三个版本的密钥管理协议：PKMvl、PKMv2、PKMv3。

PKMvl 是 IEEE 802. 16—2004 及其早前版本采用的认证与密钥管理协议，采用 X. 509 公钥证书和 RSA 算法实现了 BS 对 SS 的身份认证，进而分配授权密钥（AK）和业务加密密钥（TEK）。由于实现了 BS 对 SS 的认证，因此一定程度上阻止了非法用户接入 WiMAX 网络。但是由于仅实现了 BS 对 SS 的认证，存在伪装 BS 攻击等风险。

为应对 PKMvl 存在的诸多安全隐患，并满足由移动性带来的新安全需求，IEFE 802. 16e 标准制定了新的密钥管理协议（PKMv2）。PKMv2 首先支持 SS/MS 和 BS 之间的双向认证，同时引入了基于 EAP 的认证方法，该方法具备灵活的可扩展性，支持 EAP－AKA 和 EAP－TLS 等多种认证。此外，PKMv2 还增加了抗重放攻击措施，以及对组播密钥的管理。尽管弥补了 PKMvl 的一些安全漏洞，但 PKMv2 协议依然存在管理消息缺乏保护、DoS/DDoS 攻击和不安全的组播密钥管理三类主要安全缺陷。

PKMv3 是 IEEE 802. 16m 标准草案中新增加的一个密钥管理协议，主要目的是满足 IMT－Advanced 及实际应用环境的安全需求。与 PKMv2 相比，PKMv3 主要有两点变化：一是针对 PKMvl 和 PKMv2 存在的管理信息未加密与认证等方面的缺陷，对管理消息采取了选择性机密保护策略；二是在用户认证方法上，PKMv2 同时支持基于 EAP 认证和基于 RSA 认证两种方式，而 PKMv3 则删除了基于 RSA 认证的方式，只支持基于 EAP 认证的方式。由于 PKMv3 对管理消息采取了选择性机密保护策略，所以较好地解决了 PKMvl 和 PKMv2 存在的管理消息保护不足导致的重放攻击、DoS 攻击等安全问题。同时，由于 PKMv3 只支持 EAP 认证方式，所以增加了与其他网络交互的灵活性与兼容性。PKMvl、PKMv2、PKMv3 密钥管理协议对比见表 10-2。

表 10-2　密钥管理协议对比

安　全		PKMv1	PKMv2	PKMv3
特性	认证方式	单向	双向	双向
	认证方式	RSA	RSA 或 EAP	EAP
	管理消息保护	无	部分消息保护	分级保护
	组播	不支持	支持	支持
安全防护不足		不能阻止伪基站攻击、重放攻击、DoS 攻击，缺少管理消息保护和组播密钥管理支持	不能阻止 DoS 攻击，管理消息未保护，组播密钥管理不安全等	不能阻止 DoS 攻击

10. 2. 3　IEEE 802. 16m 安全机制

1. 标准背景

2006 年 12 月，IEEE 标准委员会通过了 IEEE 802. 16 工作组提交的 IEEE 802. 16m 标准

的立项申请。IEEE 委员会之所以决定制定 IEEE 802.16m 标准，首先是因为 IEEE 802.16e 标准在某些方面提供的能力有限，需要进一步增强；其次是来自其他标准组织的压力和挑战：3GPP 当时正在制定 LTE 标准，3GPP2 也已完成 UMB 标准的制定；最后，ITU 当时也计划于 2008 年年底至 2009 年年初正式收集 IMT－Advanced 标准候选技术。

IEEE 802.16m 项目的主要目标有两个：一是满足 IMT－Advanced 的技术要求；二是保证与 IEEE 802.16e 兼容。为了满足 IMT－Advanced 所提出的技术要求，IEEE 802.16m 下行峰值速率应该实现低速移动、热点覆盖场景下的传输速率达到 1Gbps 以上，高速移动、广域覆盖场景下的传输速率达到 100Mbps。为了兼容 IEEE 802.16e 标准，IEEE 802.16m 标准考虑在 IEEE 802.16 原有标准的基础上来实现，通过对 IEEE 802.16 原有标准的增补，进一步提高系统吞吐量和传输速率。IEEE 802.16m 标准也将兼容其他基于 OFDMA 标准的 4G 无线网络。

2009 年 7 月 30 日，IEEE 正式发布了 IEEE 802.16m 的第一个标准草案。2009 年 10 月，IEEE 802.16m 与 LTE－Advanced 一起被国际电信联盟（ITU）正式确定为 IMT－Advanced 国际标准候选技术。2011 年 4 月 1 日，IEEE 正式批准 IEEE 802.16m 成为下一代 WiMAX 标准。2012 年 1 月 18 日，ITU 正式审议通过将 WirelessMAN－Advanced（IEEE 802.16m）技术规范确立为 IMT－Advanced（4G）国际标准。

2. 体系架构

IEEE 802.16m 空中接口系统的安全功能如图 10-10 所示，主要通过对 AMS 与 ABS 之间 MAC PDU 的密码变换实现包括隐私保护、身份鉴别和机密性保护等在内的安全防护。



图 10-10　IEEE 802.16m 空中接口系统的安全功能

如图 10-10 所示，安全架构主要由两部分组成：安全管理、加密和完整性保护。

① 安全管理的功能。

- 安全管理与控制。
- EAP 加密封装/解密封装：提供与 EAP 层的接口。
- 密钥管理：负责控制所有安全组件，各种密钥在此层计算和生成。
- 授权与 SA 控制：控制认证状态机和业务加密密钥状态机。
- 位置保护：处理位置隐私相关信息。

② 加密和完整性保护的主要功能。

- 传输数据加密/认证处理：加密和解密业务数据，执行业务数据认证功能。
- 控制信息认证处理：执行消息认证功能，如 CMAC。
- 控制信息机密性保护：加密和解密控制信息，执行控制信息的认证功能。

为了满足 IMT - Advanced 及实际的应用环境的安全性需求，IEEE 802.16m 标准在用户隐私保护、管理信息安全和加密认证方式等方面较 IEEE 802.16j 等协议存在如下变化。

1）控制层面信令保护

针对 IEEE 802.16j 等之前标准版本存在的管理信息未加密与认证等方面的缺陷，IEEE 802.16m 对 MAC 控制消息提供了选择性机密性和完整性保护。IEEE 802.16m 对管理消息采取的三级选择性机密保护机制，分为基于 AES - CCM 的认证加密、基于 CMAC 的完整性保护和无保护三类，如图 10-11 所示。

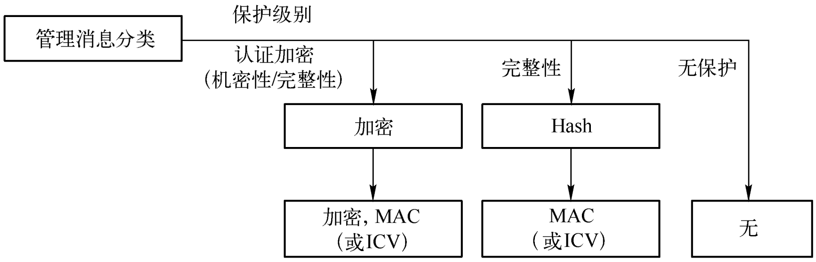


图 10-11　IEEE 802.16m 管理消息分级保护机制

- 基于 AES - CCM 的认证加密：有效载荷被加密以后，完整性检验值（ICV）被添加在其后。ICV 完整性保护对象包括有效载荷和 MAC 头两部分。
- 基于 CMAC 的完整性保护：CMAC 的完整性保护对象为整个 MAC 控制信息，仅提供完整性保护，不提供机密性保护，控制消息为明文传输。
- 无保护：如果 AMS 和 ABS 没有共享安全上下文，或者不需要提供安全保护，则管理消息既不加密，也不认证。授权阶段之前的管理消息也属于这个范畴。

2）身份信息保护

为了使 AMS 的 MAC 地址不被暴露于空中接口中，IEEE 802.16m 标准提供了 AMS 身份保护功能，通过在空中接口使用 AMSID′来减少 AMSID 的暴露时间。当选择 AMSID 保护时，AMS 在网络接入时将在 AAI - RNG - REQ 消息中采用 AMSID′作为其临时身份发送给 ABS。AMSID′为 AMSI 的 Hash 值，具体计算方式如下：

$$AMSID' = Dot16KDF(AMSID \mid 80-bit\ zero\ padding, NONCE_{AMS}, 48)$$

式中，$NONCE_{AMS}$为由 AMS 产生的一个 64bit 的随机数。

AMSID 保护的使用与否由 AMS 的 AMSID 保护策略来决定。当 S - SFH 网络配置比特为 0b0 时，网络必须允许 AMS 基于 AMSID 安全模式或利用 AMS MAC 地址（如非 AMSID 保护模式下）来接入网络；当 S - SFH 网络配置比特为 0b1 时，需要将真正的 AMSID 以明文方式发送给 ABS，因此无法进行 AMS 身份保护。

3）位置信息保护

STID 为 AMS 的唯一身份标识，在接入网络时由 ABS 分配。为了不暴露 AMSMAC 地址与 STID 之间的映射关系，IEEE 802.16m 提供了 AMS 位置信息保护功能，主要通过在初始测量过程中为 AMS 分配一个临时 STID（TSTID）来实现。TSTID 一直使用到 STID 被成功分配。具体过程如图 10-12 所示。

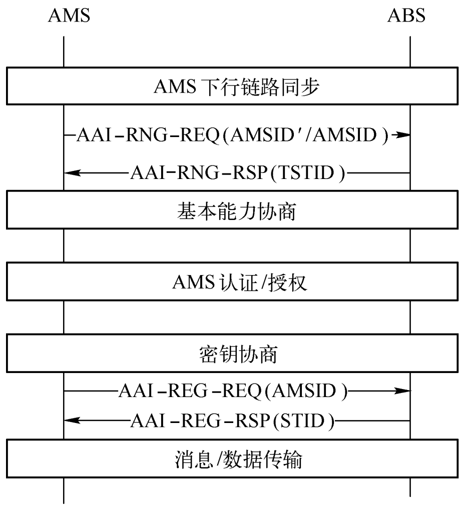


图 10-12　位置信息保护流程

AMS 产生一个新的 NONCE_AMS，而且当 S-SFH 网络配置比特为 0b0 且 AMSID 保护使能时，还要计算 AMSID′，然后将带有 AMSID′的 AAI-RNG-REQ 消息发送至 ABS。ABS 接收到 AAI-RNG-REQ 消息后，返回 AAI-RNG-RSP 消息，其中即含有 TSTID 和AMSID′。否则，当 S- SFH 网络配置比特为 0b0 且 AMSID 保护失效，或 S-SFH 网络配置比特为 0b1 时，在 AAI-RNG-REQ 和 AAI-RNG-RSP 消息中将使用 AMSID，而不是 AMSID′。

TSTID 在成功分配后，将应用于后面的网络接入过程，直到 STID 被分配。真正的 AMSID 在 AAI-REG-REQ 消息中以密态方式发送给 ABS，ABS 在收到注册请求消息后，将为 AMS 分配 STID，且加密后在 AAI-REG-RSP 消息中发送给 AMS。AMS 收到 STID 后释放 TSTID，且在后续交互中一直使用 STID 作为其身份标识。

4）密码算法选择

由于 DES 加密方式已经变得不再安全，IEEE 802.16m 标准不再支持这种加密方式，只支持 AES 加密。此外，在消息认证码方式上，IEEE 802.16j 之前的版本同时支持 CMAC 和 HMAC 两种消息码认证方式，但 IEEE 802.16m 只支持 CMAC 消息认证方式。

5）认证方式选择

在用户认证方法上，IEEE 802.16j 之前的标准版本同时支持基于 EAP 认证和基于 RSA 认证两种认证方式。考虑到基于 EAP 认证的方法与 AAA 体系结构交互的灵活性，在 IEEE 802.16m 标准中删除了基于 RSA 认证的方式，只支持基于 EAP 认证的方式。同时，在密钥协商过程中引入了随机数 nonce 和计数器 count 等新的密钥参数。

3. 密钥管理协议（PKMv3）

IEEE 802.16m 使用 PKMv3 实现以下功能：认证与授权消息透明交换、密钥协商、安全材料交换。PKMv3 提供 AMS 与 ABS 之间的双向认证，并且通过认证建立双方之间的共享密钥，利用共享密钥实现其他密钥的交换与派生。这种机制可以在不增加运算操作的基础上，实现业务密钥的频繁更换。

IEEE 802.16m 标准继承了 IEEE 802.16j 的密钥体系，主要密钥包括主密钥（MSK）、PMK、AK、TEK 和 CMAC keys 等，具体层次关系如图 10-13 所示。

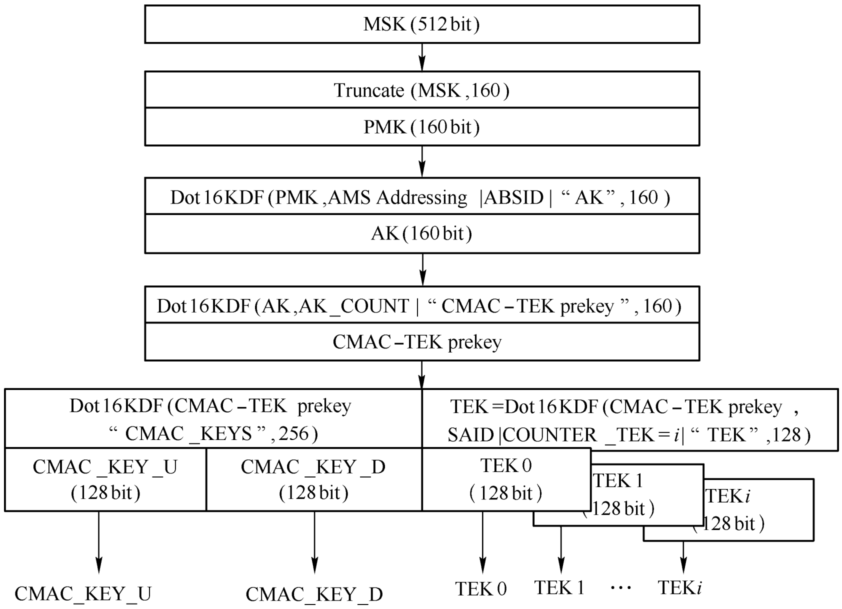


图 10-13　IEEE 802.16m 密钥体系

AMS 在申请接入 IEEE 802.16m 网络时，首先和 ABS 利用 EAP 认证方法进行双向认证，认证通过后，它们将共享一个 512bit 的主密钥 MSK，随后 AMS 和 ABS 通过三次握手协议交换密钥资源，并利用事先协商好的密钥生成算法陆续获取 PMK、AK、TEK 和 CMAC keys 等密钥。

TEK 是通信数据加密密钥，它又被分为上行链路 UL（Uplink）加密密钥 TEK_U 和下行链路 DL（Downlink）加密密钥 TEK_D。CMAC keys 是用于管理消息认证的 CMAC 码密钥，它也被分为上行链路 CMAC 码密钥 CMAC_KEY_U 和下行链路 CMAC 码密钥 CMAC_KEY_D。PMK、AK、TEK 和 CMAC keys 的具体产生方法分别定义如下：

$$\text{PMK} = \text{truncate}(\text{MSK}, 160)$$

$$\text{AK} = \text{Dot16KDF}(\text{PMK}, \text{MS Addressing} \mid \text{ABSID} \mid \text{AK}, 160)$$

$$\text{CMAC} - \text{TEK prekey} = \text{Dot16KDF}(\text{AK}, \text{AK_COUNT} \mid \text{CMAC} - \text{TEK prekey}, 160)$$

$$\text{CMAC_KEY_U} \mid \text{CMAC_KEY_D} = \text{Dot16KDF}(\text{CMAC} - \text{TEK prekey}, \text{CMAC_KEYS}, 256)$$

$$\text{TEKi} = \text{Dot16KDF}(\text{CMAC} - \text{TEK prekey}, \text{SAID} \mid \text{COUNTER}_{\text{TEK}} = i \mid \text{TEK}, 128)$$

其中，truncate(x,y)表示 x 的后 y bit 数据，“Dot16KDF” 是 IEEE 802.16m 标准中定义的密钥分发函数，“TEK Counter” 表示针对相同 SAID 产生不同 TEKs 的计数器，“|” 表示连接符。

Dot16KDF 函数的定义如下：

```
Dot16KDF(key,astring,keylength)
  {
result = null;
```

```
        Kin = Truncate(key,128);
        for(i = 0;i <= int((keylength - 1)/128);1 ++)
        {
            result = result | CMAC(Kin,i | astring | keylength);
        }
    return Truncate(result,keylength);
    }
```

具体密钥协商流程如图 10-14 所示。

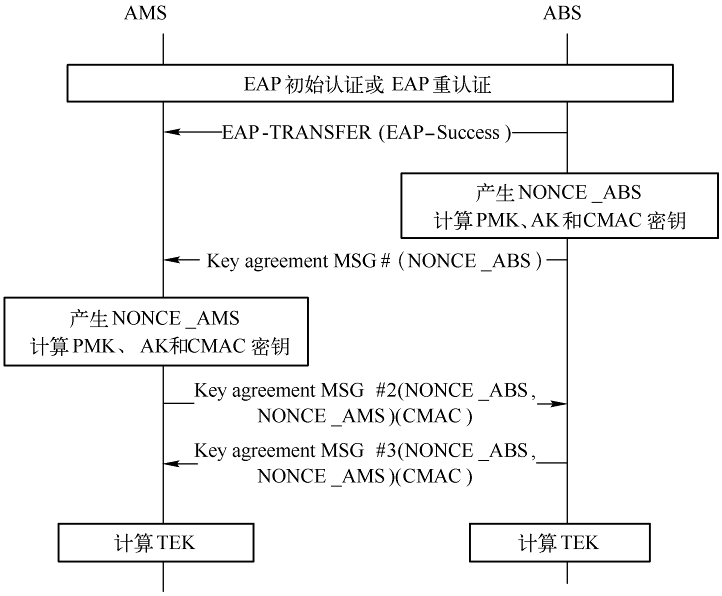


图 10-14　密钥协商流程

步骤一：ABS 首先产生一个随机数 NONCE_ABS，然后给 AMS 发送密钥协商询问信息（key agreement MSG#1）。

步骤二：AMS 收到 Key agreement MSG#1 后，首先检查 NONCE_ABS 的新鲜性，如果不是新鲜的，则丢弃 Key agreement MSG#1；否则，AMS 产生一个随机数 NONCE_AMS，并利用 MSK、NONCE_BS、NONCE_MS 和 MSID′（AMS 的 MAC 地址的变换）等密钥资源生成 PMK、AK 和 CMAC keys，最后 AMS 利用生成的 CMAC keys 将 MSID′、NONCE_BS、NONCE_MS 等生成 CMAC 值，并给 ABS 发送密钥协商请求消息（Key agreement MSG#2）。

步骤三：接收到 Key agreement MSG#2 后，ABS 采取与 AMS 协商好的方式生成 PMK、AK 和 CMAC keys，并利用 CMAC keys 验证 Key agreement MSG#2 中的 CMAC 值，如果验证正确，则利用 CMAC keys 将 NONCE_BS、NONCE_MS 等生成 CMAC 值，并给 AMS 发送密钥协商响应消息（Key agreement MSG#3）。

步骤四：接收到 Key agreement MSG#3 后，AMS 利用之前生成的 CMAC keys 验证 Key agreement MSG#3 中的 CMAC 值，如果验证正确，则为支持的 SA 生成 TEK。

10.3　3G 和 LTE 安全技术

10.3.1　3G 移动通信网络及安全威胁

1. 3G 通信网络组成

第三代移动通信包括 WCDMA、CDMA2000 和 TD-SCDMA 三种标准。在国际电信联盟的 3G 框架中，主要推广的还是 UMTS。UMTS 融合了 TDMA、CDMA 的关键技术和集成的卫星组件，在移动通信网络中提供宽带多媒体业务。在这里，对 3G 系统的研究主要是指以 UMTS 为例的 3G 系统的研究。

UMTS 的系统构架如图 10-15 所示。从功能上对 UMTS 系统分组，其网络单元可以分为 RAN 和 CN 两部分。RAN 的主要功能是处理所有与无线相关的业务；CN 的主要功能是处理 UMTS 系统内的所有语音呼叫和数据连接业务，并实现与外部网络间的交换和路由功能。整个 UMTS 系统是由 RAN、CN 和 UE（用户设备）构成的。

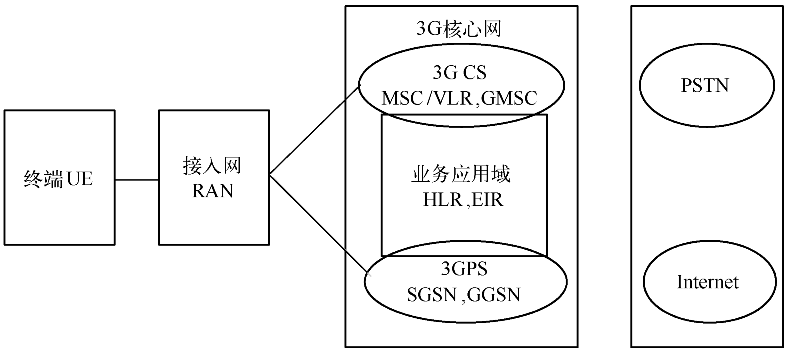


图 10-15　UTMS 的系统架构

2. 3G 系统面临的安全威胁

① 对敏感数据的非法获取，对系统信息的保密性进行攻击。

- 窃听：攻击者对通信链路进行非法窃听，获取消息。
- 伪装：攻击者伪装合法身份，诱使用户或网络相信其身份合法，从而窃取系统信息。
- 业务分析：攻击者对链路中信息的时间、速率、长度、消息源及目的地等信息进行分析，从而判断用户位置或了解正在进行的重要的商业交易。
- 浏览：攻击者对敏感信息的存储位置进行搜索。
- 泄露：攻击者利用合法接入进程获取敏感信息。
- 试探：攻击者通过向系统发送信号来观察系统反应。

② 对敏感数据的非法操作，对信息的完整性进行攻击，攻击者故意对信息进行篡改、插入、重放或删除。

③ 对网络服务干扰或滥用，从而导致系统拒绝服务或导致系统服务质量的降低。

- 干扰：攻击者通过阻塞用户业务、信令或控制数据使合法用户无法使用网络资源。

- 误用权限：用户或服务网络可以利用它们的权限来越权获得业务或信息。
- 服务滥用：攻击者通过滥用某些特定的系统服务，从而获取好处，或者导致系统崩溃。
- 拒绝：用户或网络拒绝发出响应。

④ 否认。主要指用户或网络否认曾经发生的动作。

3. 针对 3G 系统的攻击方法

针对 3G 的攻击方法主要包含针对系统无线接口的攻击、针对系统核心网的攻击和针对终端的攻击三种方式。

1）针对系统无线接口的攻击

① 对非授权数据的非法获取：基本手段主要包括对用户业务的窃听、对信令与控制数据的窃听、伪装网络实体截取用户信息，以及对用户流量进行主动与被动分析。

② 对数据完整性的攻击：基本手段主要包括对系统无线链路中传输的业务与信令、控制消息进行篡改，包括插入、修改、删除等。

③ 拒绝服务攻击。

④ 对业务的非法访问攻击：攻击者伪装其他合法用户身份，非法访问网络，或切入用户与网络之间，进行中间人攻击。

⑤ 用户身份捕获攻击：攻击者伪装成服务网络，对目标用户发起身份请求，从而捕获用户明文形式的永久身份信息。

⑥ 对目标用户与服务网络之间的加密流程进行压制，使加密流程失效。

2）针对系统核心网的攻击

① 对数据的非法获取：基本手段包括对用户业务、信令及控制数据的窃听，冒充网络实体截取用户业务及信令数据，对业务流量的被动分析，对系统数据存储实体的非法访问，以及在呼叫建立阶段伪装用户位置信息等。

② 对数据完整性的攻击：基本手段包括对用户业务与信令消息进行篡改，对下载到用户终端或 USIM 的应用程序及数据进行篡改，通过伪装成应用程序及数据的发起方篡改用户终端或 USIM，篡改系统存储实体中存储的用户数据等。

③ 拒绝服务攻击：基本手段包括物理干扰、协议干扰、伪装成网络实体对用户请求做出拒绝回答等。

④ 否定：主要包括对费用的否定、对发送数据的否定、对接收数据的否定等。

⑤ 对非授权业务的非法访问：基本手段包括伪装成用户、服务网络、归属网络，滥用特权非法访问非授权业务。

3）针对终端的攻击

主要是针对终端和 USIM 的攻击，包括使用偷窃的终端和 USIM 对终端或 USIM 中数据进行篡改，对终端与 USIM 间的通信进行窃听，伪装身份截取终端与 USIM 间的交互信息，非法获取终端或 USIM 中存储的数据。

10.3.2　3GPP 安全增强技术

3G 系统的安全问题在其标准制定之初已经进行了考虑，3G 网络结构采取由第二代移动

通信系统演进的策略，因此其安全机制保留了 GSM 或其他第二代移动通信系统的优良安全策略，并改进了其中的诸多不足。同时，3G 安全体制还对 3G 中出现的新业务提供了安全保护。这些安全措施使 3G 系统的安全性能有了很大提高。

3G 系统安全逻辑结构分为三层，定义了五组安全特性。

① 网络接入安全。主要针对无线链路的攻击，包括用户身份保密、用户位置保密、用户行踪保密、实体身份认证、加密密钥分发、用户数据与信令数据的保密及消息认证。

② 网络域安全。主要保证核心网络实体间安全交换数据，包括网络实体间身份认证、数据加密、消息认证及对欺骗信息的收集。

③ 用户域安全。主要保证对移动台的安全接入，包括用户与智能卡间的认证、智能卡与终端间的认证及链路的保护。

④ 应用域安全。用来在用户和服务提供商应用程序间提供安全交换信息的一组安全特征，主要包括应用实体间的身份认证、应用数据重放攻击的检测、应用数据完整性保护、接收确认等。

⑤ 安全特性可见性及可配置能力。主要指用户能获知安全特性是否在使用，以及服务提供商提供的服务是否需要以安全服务为基础。

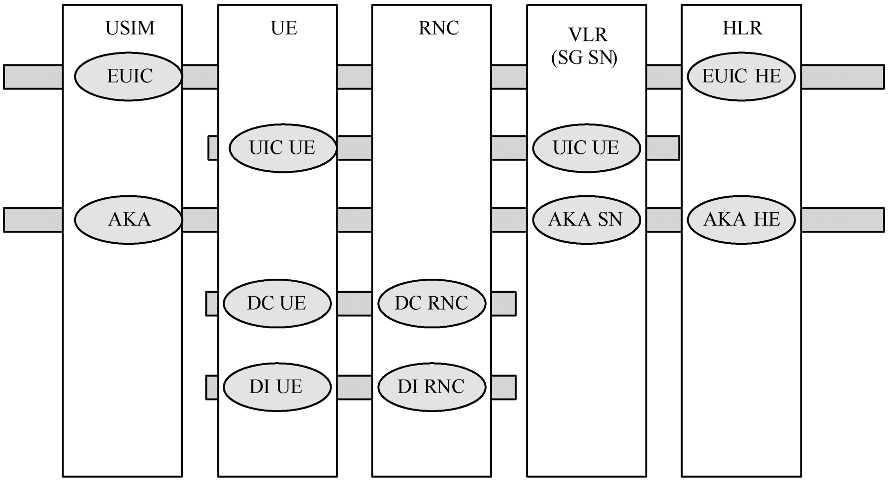
由于在第三代移动通信系统中，终端设备和服务网间的接口是最容易被攻击的点，所以如何实现更加可靠的网络接入安全能力，是 3G 系统安全方案中至关重要的一个问题。网络安全接入机制应该包括用户身份保密、接入链路数据的保密性和完整性保护机制，以及认证和密钥分配机制。

3G 安全功能结构如图 10-16 所示。其中，横轴代表网络实体，涉及的网络实体依据利益关系分为三部分：用户部分，包括用户智能卡（USIM）及用户终端（UE）；服务网络部分，包括服务网络无线接入控制器（RNC）和拜访位置寄存器（VLR）；归属网络部分，包括用户位置寄存器（HLR）和认证中心（UIDN，图中未显示）。纵轴代表相应的安全措施，主要分为：增强用户身份保密（EUIC），通过归属网内的 UIDN 对移动用户智能卡身份信息进行认证；用户与服务网间身份认证（UIC）；认证与密钥协商（AKA），用于 USIM、VLR、HLR 间的双向认证及密钥分发；用户及信令数据保密（DC），用于 UE 与 RNC 间信息的加密；消息认证（DI），用于对交互消息的完整性、时效及源与目的地进行认证。

相对于 GSM 网络，3GPP 网络主要进行了如下改进。

- 实现了双向认证：不但提供网络对 MS 的认证，也提供了 MS 对网络的认证，可有效防止伪基站攻击。
- 密钥长度增加为 128bit，改进了算法。
- 3GPP 接入链路数据加密延伸至无线接入控制器 RNC，提供了接入链路信令数据的完整性和加密保护。
- 3G 的安全机制还具有可拓展性，为将来引入新业务提供了安全保护措施。
- 3G 能向用户提供安全可视性操作，用户可随时查看自己所用的安全模式及安全级别。

在密钥长度、算法选定、鉴别机制和数据完整性检验等方面，3GPP 网络的安全性能都远远优于 2G 网络。



USIM：用户业务识别模块　DI：数据完整性　SGSN：服务GPRS支持节点
EUIC：增强用户身份保密　DC：数据加密　SN：服务网络
AKA：认证与密钥分配　RNC：无限网络控制器　HLR：归属位置寄存器
UE：用户终端　VLR：访问位置寄存器
UIC：用户与服务网身份认证　HE：本地环境

图 10-16　3G 安全功能结构

10.3.3　LTE/SAE（4G）安全技术

1. LTE 系统架构

为了进一步发展 3G 技术，3GPP 于 2004 年将 LTE（Long Time Evolution）作为 3G 系统的长期演进。在开展 LTE 研究项目的同时，启动了 SAE（System Architecture Evolution）的研究项目。LTE 体系结构可以借助 SAE 体系结构来做详细描述。LTE 致力于无线接入网的演进（E-UTRAN），SAE 则致力于分组网络的演进（EPC）；LTE 和 SAE 共同组成演进型分组系统（EPS）。EPS 的系统架构如图 10-17 所示。

EPS 系统架构可以划分为 4 个域，即 UE、演进 UTRAN（E-UTRAN）、EPC 和业务域。UE、E-UTRAN 和 EPC 共同组成 IP 连接层，这部分也称演进分组系统（EPS），该层的主要功能是提供基于 IP 的连接。所有业务将在 IP 连接之上提供，且在 E-UTRAN 和 EPC 中不存在早期 3GPP 架构所包含的电路交换节点和接口。

IP 多媒体子系统（IMS）是一个典型的业务设备实例，该设备可在业务连接层中使用，以在低层 IP 连接之上提供各种业务。例如，为支持语音业务，IMS 可以提供 IP 语音（VoIP），并可以通过由它控制的网关实现与传统电路交换网络 PSTN 和 ISDN（综合业务数字网）的互连。

E-UTRAN 系统由 eNode B（又可简写为 eNB）组成，提供用户层面和控制层面（RRC）协议，用户层面包括分组数据汇聚（PDCP）协议、无线链路控制（RLC）协议、介质访问控制（MAC）协议和物理层（Physical layer）协议，控制层面包括无线资源控制（RRC）协议。eNB 之间通过 X2 接口互连，同时 eNB 也通过 S1 接口和 EPC 相连。更细地划

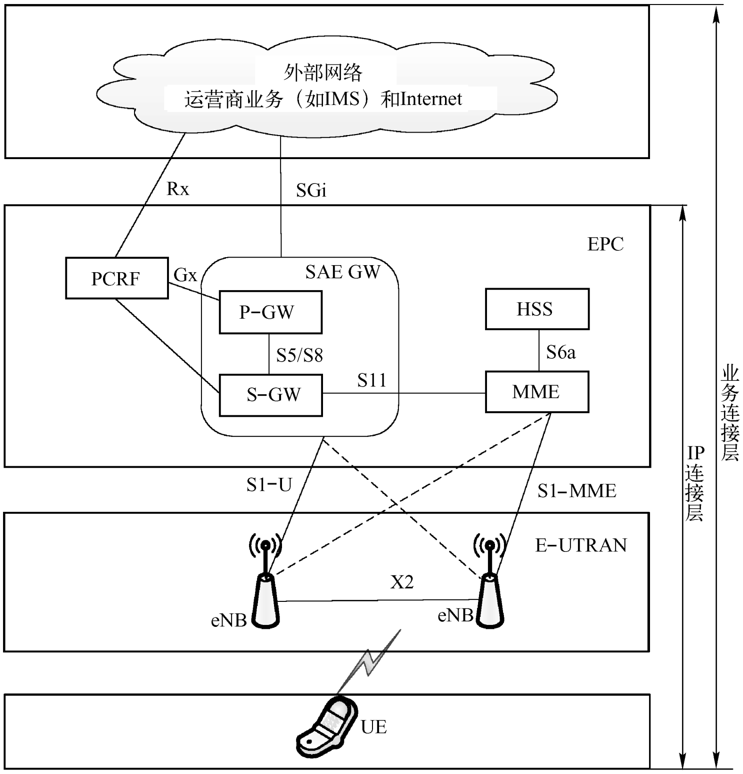


图 10-17　EPS 的系统架构

分，即通过 Sl－MME 接口与移动管理实体（MME）互连，通过 Sl－U 接口与服务网关（S－GW）互连。Sl 接口支持多个 eNB 与多个 MME 和 S－GW 互连。

E－UTRAN 系统的总体架构如图 10-18 所示。

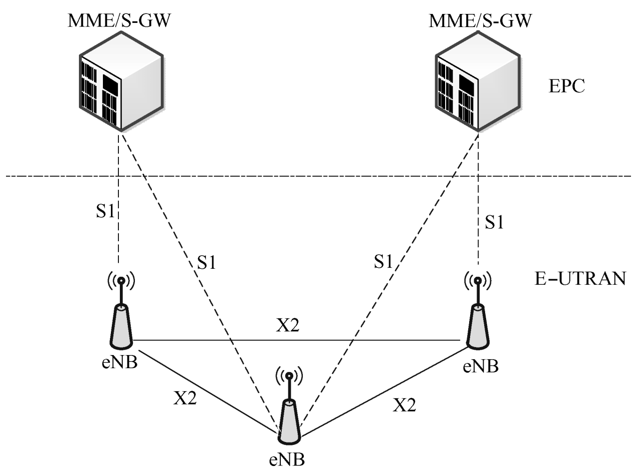


图 10-18　E－UTRAN 系统的总体架构

eNB 的主要功能描述如下。

① 无线资源管理功能。包括无线承载控制、无线接纳控制、连接移动性控制、上行和

下行资源动态分配（调度）。

② IP 头压缩和用户数据流加密。

③ 当根据 UE 提供的信息无法获取 MME 路由时，选择为附着 UE 选择 MME。

④ 将用户层面数据路由至服务网关。

⑤ 调度和传输寻呼信息。

⑥ 调度和传输广播信息。

⑦ 移动和调度的策略和测量报告配置。

MME 的主要功能包括以下几个方面。

① 非接入层（NAS）信令。

② NAS 信令安全。

③ 接入层（AS）安全控制。

④ 漫游管理。

⑤ 认证鉴别。

⑥ 承载控制。

各个实体的功能如图 10-19 所示。

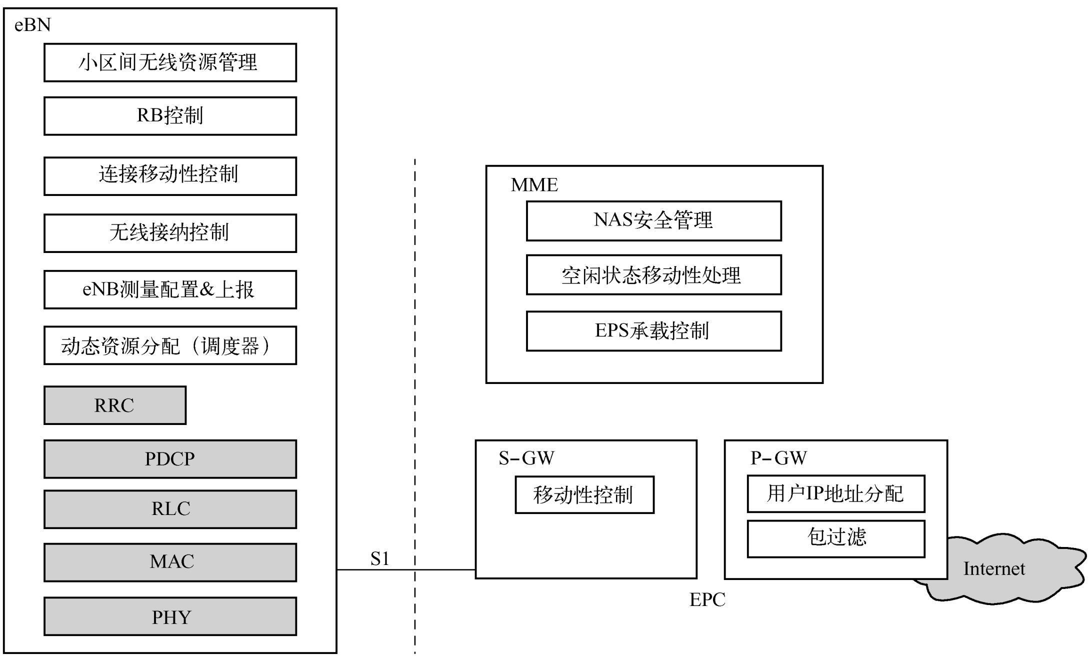


图 10-19　E－UTRAN 和 EPC 的功能划分

2. LTE 安全体系架构

LTE/SAE 安全体系架构如图 10-20 所示，主要包括网络接入安全、网络域安全、用户域安全、应用域安全、安全可视性与可配置性五个安全特性组，每个安全特性组可以应对不同的安全威胁，完成特定的安全性目标。

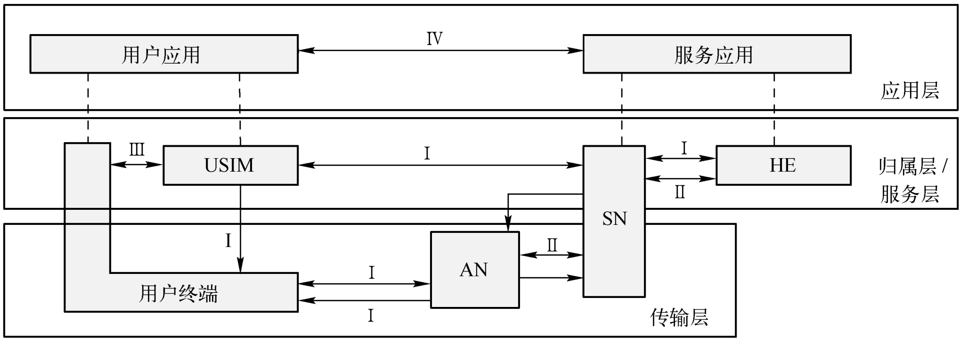


图 10-20　安全体系架构

① 网络接入安全（Ⅰ）：为用户接入服务网络提供安全防护，主要解决无线接入通信链路安全问题。

② 网络域安全（Ⅱ）：为节点间信令数据、用户数据安全交换提供保证，应用于 AN 和 SN 之间及 AN 之间，主要用于防止有线网络攻击。

③ 用户域安全（Ⅲ）：主要为移动台提供接入安全防护。

④ 应用域安全（Ⅳ）：为用户域和服务域之间的各种应用提供安全数据交换保证。

⑤ 安全的可视性和可配置性：该安全特性组能够使用户知道正在使用的安全特性，以及服务的提供和使用是否应当依赖于安全特性。

3. LTE 安全特性与需求

目前，3G 系统安全防护功能已经得到广泛的应用，可以为 3G 网络提供用户 ID 机密性保护、双向认证鉴别、用户层面（U - Plane）机密性保护、控制层面（C - Plane）机密性和完整性保护等功能。LTE 架构设计与 3G 系统有较大的区别，为适应这些变化，需要改进相应的安全功能。LTE 安全功能总体需求如下。

- 在不影响用户应用方式的前提下，提供至少与 3G 同等安全级别的防护。
- 能够抵御现有 Internet 攻击。
- LTE 提供的安全功能不能影响由 3G 向 LTE 演进。
- 允许继续使用用户身份模块（USIM）。

LTE 具体安全功能需求分析如下。

1）用户到网络的安全

① 用户身份机密性。LTE 用户身份机密性需求与 3G 相同，其安全特征及需求如下。

- 用户身份机密性：用户身份（IMSI）不能在无线链路中被窃听。
- 用户位置机密性：在某一区域中，用户的存在和到达不能通过无线链路侦听被判断。
- 用户不可追踪性：窃听者不能通过无线链路侦听来推断是否不同的服务被提供给同一个用户。

用户身份机密性可以通过为用户终端分配临时身份标识（TMSI）来实现，为避免被跟踪，用户终端不应长期使用同一个临时身份标识。另外，为了实现这些安全特性，在无线接入链路中，所有可能暴露用户身份的任何信令或用户数据都应被加密保护。

② 设备保密性。从用户隐私的角度考虑，MSIN（移动用户识别号码）、IMEI（国际移

动设备身份码）和 IMEISV（国际移动设备身份码版本号）应被加密保护。

IMEI 和 IMEISV 应该被安全地存储在终端中。在被完整性保护的请求中，UE 应在网络请求时提供它的设备标识符 IMEI 或 IMEISV 给网络。在 NAS 协议中，若 NAS 安全未被激活，UE 则不给网络发送 IMEI 和 IMEISV 来响应一个网络请求。但是，在紧急连接期间，当 UE 没有有效 IMSI、GUTI 或 P-TMSI 时，在 NAS 安全激活之前，UE 响应网络请求的消息中应该包括 IMEI。在某些情况下（例如，在第一次连接过程中），MSIN 必须以明文形式发送给网络。当 NAS 机密性保护超过运营商的选择范围之外时，IMEI 和 IMEISV 则不能被机密性保护。

③ 实体认证。实体认证与 3G 中所描述的实体认证相同。实体认证的安全特性包括以下两个方面。

- 用户认证：服务网络对用户身份的认证。
- 网络认证：用户对服务网络的认证，使得用户能够证实其所连接的服务网络已被授权，且该授权是最新的。

实体认证应该发生在用户与网络的每个连接建立期间。为了取得实体双向认证，LTE 中也采用了和 3G 相同的两种机制：一是使用由用户 HE（归属域）传递到服务网络的认证向量来实现认证；二是本地认证机制，即在认证与密钥协商期间，通过使用在用户和服务网络间建立的完整性密钥来实现认证。

④ 数据机密性与完整性。

- 用户数据和信令数据的机密性。

LTE 系统可以提供 RRC 信令和 NAS 信令机密性保护功能，这些功能都是可配置的可选项，通过 RRC 信令加密，可以有效阻止攻击者通过小区级测量报告、切换消息映射、小区级身份链等信息跟踪特定用户。

在受限制服务模式下的紧急呼叫期间，当 UICC 上的证书认证不能成功执行时，RRC 和 NAS 信令、用户层数据的机密性保护应被省略。此时，应该由网络选择 EEAO（NULL 密码算法）来为 NAS、RRC 和用户层数据提供机密性保护。

- 用户数据和信令数据的完整性。

为了防止信令被篡改，LTE 要求 NAS 信令和 RRC 信令都应该具备完整性和抗重放保护措施。但考虑到用户数据通信的低延迟特性，LTE 系统在 eNB 和 UE 之间的用户层数据分组不要求提供完整性保护。

在受限制服务模式下的紧急呼叫期间，当 UICC 上的证书认证不能被成功地执行时，RRC 和 NAS 信令的完整性保护和重放保护应被省略。此时，应由网络选择 EIA0（NULL 密码算法）来为 NAS、RRC 提供完整性保护。

2）安全可见性和可配置性

① 安全可见性。尽管一般而言，安全特性对于用户来说应该是透明的，但对于某些特定事件和具体用户关心的内容，应该向用户提供以下安全操作的可见性：接入网络加密的指示时，用户应被告知无线接入链路中是否采用了用户数据机密性保护机制，尤其是当前在非加密呼叫模式的情况下。

② 可配置性。可配置性是指用户能够配置服务的使用情况或配置特定的安全特性，仅当所有用户所配置的相关安全特性被实施时，该服务才能被使用。建议使用的可配置特性包

括启用/禁用用户的 USIM 认证（用户应该能够控制 USIM 认证操作）。

3）关于 eNB 的安全要求

① eNB 的建立和配置需求。建立和配置 eNB 应该被认证和授权，以使得攻击者不能够通过本地和远程接入来修改 eNB 的设置和软件配置。

- 在 EPS 核心网和 eNB 之间，以及通过 X2 连接的相邻 eNB 之间，都需要建立安全关联。这些安全关联的建立都应该进行相互认证，并用于实体间的通信。
- 远程/本地 O&M（操作和管理）系统与 eNB 之间的通信应相互验证。
- eNB 应能确保软件/数据更改已被授权。
- eNB 应使用被授权的数据/软件。
- 对启动过程的敏感部分，应在安全环境下执行。
- 应确保发往 eNB 软件的机密性。
- 应确保发往 eNB 软件的完整性。

② eNB 内部的密钥管理需求。EPS 核心网络为 eNBs 提供用户特定的会话密钥资源，它同时也拥有用于认证和安全关联建立的长期密钥。存储在 eNBs 内的密钥应不能离开 eNB 内的安全环境。

③ eNB 用户层面数据的处理需求。eNB 主要负责在 Uu 参考点和 Sl/X2 参考点之间的用户层数据加解密。

- 用户层面数据加密/解密应在安全环境中完成，相关密钥也在此安全环境中存储。
- Sl－U 和 X2－U 上传输的用户数据应提供完整性、机密性和抗重放保护，阻止非授权用户的攻击。

④ eNB 控制面数据的处理需求。eNB 需要为 Sl/X2 上传输的控制层面数据提供机密性和完整性保护。

- 控制层面数据加密/解密应在安全环境中完成，相关密钥也在此安全环境中存储。
- Sl－MME 和 X2－C 上传输的控制层面数据需要提供完整性、机密性和抗重放保护，防止非授权用户的攻击。

⑤ eNB 安全环境的需求。安全环境在逻辑上定义在 eNB 内部，由支持各种敏感操作的多种功能构成。

- 安全环境应该支持敏感数据的安全存储，如长期密码资源和重要的配置数据。
- 安全环境应该支持敏感功能的执行，如用户数据的加密/解密及密码协议处理流程（如身份认证协议）。
- 安全环境中使用的敏感数据不能暴露给外部实体。
- 安全环境应该支持启动过程中敏感组成部分的执行。
- 应该确保安全环境的完整性。
- 只有授权操作才允许接入安全环境，包括数据存储、使用及函数执行等。

LTE 系统安全技术相对 3G 系统主要有以下变化。

① 引入分级密钥架构，可以为不同的目的产生不同的密钥。

② 区分 NAS 和 AS 安全功能，NAS 安全主要应用于核心网络节点和用户终端节点，AS 主要应用于 eNB 和用户终端节点。

③ 引入前向安全概念，在部分密钥失控时可以减少危害范围。

④ 增加 3G 网络和 LTE 网络之间互连的安全功能。

参考文献

[1] 朱建明，马建峰. 无线局域网安全：方法与技术［M］. 北京：机械工业出版社，2009.
[2] 王顺满，等. 无限局域网络技术与安全［M］. 北京：机械工业出版社，2005.
[3] 王兵. WAPI 安全机制浅析［J］. 计算机安全，2011.
[4] 郎为民，刘波. WiMAX 技术原理与应用［M］. 北京：机械工业出版社，2008.
[5] 刘波，安娜，黄旭林. WiMAX 技术与应用详解［M］. 北京：人民邮电出版社，2007.
[6] 朱晓妍，田海博. TKIP 协议简介及其安全分析［M］. 西安：西安电子科技大学出版社. 2004.
[7] 刘东苏. 移动通信系统中的若干安全问题研究［D］. 西安：西安电子科技大学，2006.
[8] ETSI. ETSI TS 100 585. Version7. 0. 1 CSM Technical Specification［S］. France，ETSI Press，1999.
[9] 何红永，唐晓梅. 数字移动通信中的用户鉴权［J］. 移动通信，2001，25（7）：16 －18.
[10] Barkan Elad，Biham Eli，Keller Nathan. Instant Ciphertext – Only Cryptanalysis of GSM Encrypted Communi – cation［OL］. http：//www. cs. technion. ac. il/users/wwwb/cgi – bin/tr – get. cgi/2003/CS/CS – 200 – 05. ps. gz，2003.
[11] 赵旸，移动通信系统的安全性研究［J］. 通信与信息技术，2002（06）：54 –56.
[12] 3CPP. 3CPP TS 33. 102 3C Secunty architecture［S］. France：3GPP，2004.
[13] 谢烨. 3G 的安全体系结构简述［J］. 信息安全与通信保密，2005（3）：161 –163.
[14] 李世鸿，李方伟. 3G 移动通信中的安全改进［J］. 重庆邮电学院学报，2002，14（4）：24 –27.
[15] 3GPP. 3GPP TS 36. 300 Evolved Universal Terrestrial Radio Access（E – UTRA）and Evolved Universal Ter – restrial Radio Access Nerwork（E – UTRAN）Overall description Stage2［S］. France：3CPP，2010.
[16] 3CPP. 3CPP TS 33. 401. 3GPP System Architecture Evolution（SAE）：Security architecture［S］. France：3GPP，2009.

第 11 章　物联网核心网络安全技术

物联网核心网络主要指的就是计算机网络，相对于感知层安全而言，计算机网络的攻击方法多样，安全防护技术也相对比较成熟。总体而言，计算机网络安全防御技术主要分为被动防御和主动防御两方面，本章在介绍典型被动防御技术，如病毒检测、防火墙等安全技术的同时，还将介绍目前典型的网络主动防御技术，如入侵检测、态势感知、移动目标防御等。

11.1　被动防御——计算机病毒检测技术

11.1.1　计算机病毒

计算机病毒（Computer Virus）在《中华人民共和国计算机信息系统安全保护条例》中被明确定义为“编制或者在计算机程序中插入的破坏计算机功能或者破坏数据，影响计算机使用并且能够自我复制的一组计算机指令或者程序代码”。而在一般教科书及通用资料中，它被定义为：利用计算机软件与硬件的缺陷，由被感染机内部发出的破坏计算机数据并影响计算机正常工作的一组指令集或程序代码 。计算机病毒最早出现在 20 世纪 70 年代 David Gerrold 的科幻小说 *When H. A. R. L. I. E. was One* 中。其最早的科学定义出现在 1983 年 Fred Cohen（南加利福尼亚大学）的博士论文《计算机病毒实验》中，他指出病毒是“一种能把自己（或经演变）注入其他程序的计算机程序”。病毒的名称主要来自生物病毒，生物病毒就是把自己注入细胞之中，因此计算机病毒的名称也来源于此。

计算机病毒不是来源于突发或偶然的原因。一次突发的停电和偶然的错误，会在计算机的磁盘和内存中产生一些乱码和随机指令，但这些代码是无序和混乱的，病毒则是一种比较完美的精巧严谨的代码，按照严格的秩序组织起来，与所在的系统网络环境相适应和配合；病毒不会偶然形成，并且需要有一定的长度，这个基本的长度从概率上来讲是不可能通过随机代码产生的。图 11–1 描述了计算机病毒程序结构的基本模式。

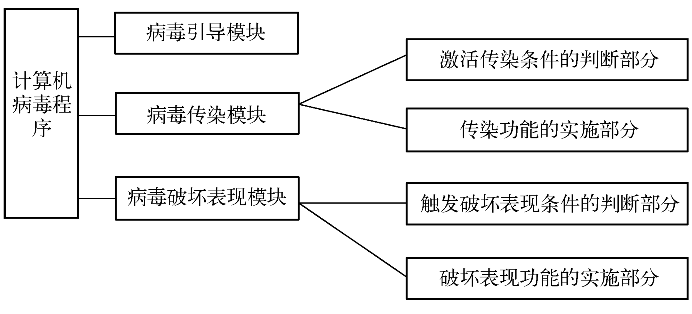


图 11–1　计算机病毒程序结构的基本模式

现在流行的计算机病毒一般都是人为故意编写的，有些是个人目的，也有些商业目的，还有些是政治目的或军事目的。例如，近年来日益增多的 APT（Advanced Persistent Threat）攻击往往是由一些特定组织发起的，他们利用先进的攻击手段对特定目标进行长期持续性网络攻击，破坏性强，防御困难，典型的 APT 攻击是美军对伊朗核电站实施的“震网”病毒。

11.1.2　计算机病毒的特点及分类

1. 计算机病毒的特点

计算机病毒其实是计算机程序的一种，在某些时候计算机恶意程序与正常用户程序几乎相同，但很多情况下，两者之间具有很大的差别。计算机病毒所具备的一些特性是其他正常程序不具备的。

1）传染性

计算机病毒最基本的特征是具有传染性。只要病毒进入系统，便会对计算机系统展开搜索以查找出满足自己传染条件的其他程序，或者选择某些文件将病毒程序插入进去，以使得自身得以不断繁殖。网络中有一台计算机被病毒感染后，如果不立即采取有效措施查杀病毒，就会造成病毒在该计算机上肆意感染其他文件。被病毒感染后的文件又会成为新的病毒感染源，且处在互联网环境中的计算机会利用所有可用的媒介进行病毒传播，从而造成整个网络的瘫痪。

2）隐蔽性

计算机病毒通常利用较高的编程技巧进行编制且具有短小精悍的特点，一般潜伏在一些比较隐蔽的地方，但不排除一些病毒通过隐藏程序的形式出现。通常情况下，当计算机系统被病毒程序感染后，用户不会感觉到病毒的存在，而只有在病毒程序被激发后造成计算机系统反应不正常时用户才会发觉已经感染病毒。例如，著名的 Tiny 家族 Dos 病毒，这类病毒都非常短小，最短的病毒程序长度只有 133B。

3）潜伏性

大多数病毒具备潜伏性，它们在感染计算机程序或文档后可能不会立即发作，而是先潜伏在计算机系统中，只有当使它们具有破坏能力的某些特殊条件被满足时才会被激发。通过这样隐藏自身，它们就可以被广泛传播。病毒的潜伏性越好，它们在计算机中隐藏的时间就会越长，从而传染范围就会越大。当病毒处在潜伏期时，如果不使用专门的计算机病毒检测程序进行病毒检测，是无法发现的。这样，病毒就可以长期潜伏在硬盘里，只要条件满足就会启动其破坏功能。

4）破坏性

计算机病毒只要进入要攻击的系统，便会对计算机系统或应用程序造成程度不同的影响或破坏。影响程度较轻时会通过占用用户资源，如系统资源等来降低计算机的正常工作效率；而重者则会造成系统的崩溃。因此，根据病毒破坏程度不同，将计算机病毒分为良性的和恶性的。某些恶性病毒可能使用户数据遭到破坏，甚至损坏系统硬件。

5）不可预见性

目前所有的计算机病毒检测方法均只能对付网络中出现过的已知病毒或极少部分未知病

毒。由于所有的计算机病毒程序检测算法复杂度都达到了 NP，因此无论哪种病毒检测方法都不可能对所有的计算机病毒均有效。病毒和病毒检测技术一直呈相互激励的提高态势，这是计算机安全领域不可否认的规律。

2. 计算机病毒的分类

计算机病毒种类多样，破坏方式也各有千秋。因此，其分类方式也要按病毒的特点进行区分。总的来说，按照连接方法、传染原理、破坏程度这三种方式进行分类是当前主流的分类方法。下面具体介绍这三种分类方法。

1）按连接方法分类

按连接方法分类就是按病毒在宿主程序中的位置进行分类，主要分为外壳型、嵌入式、源码型、操作系统型四种。外壳型病毒会寄居在宿主程序执行之前或之后，它不会主动地破坏宿主程序的代码结构，而是先于宿主程序执行，通常还会修改宿主程序的第一条指令，这样会导致病毒在宿主程序执行过程中被大范围传播，甚至难以控制。文件型病毒就是按照这种方式传播的，如已经存在的炸弹病毒（Chinese Bomb）、Jerusalem 病毒等。与外壳型病毒不同的是，嵌入式病毒寄居在宿主程序的内部，将自身嵌入宿主程序的代码段中，而且很难清除，目前大多数超级病毒都采用这种技术。如何快速高效地检测出嵌入式病毒是目前计算机安全领域的重大挑战。该病毒诞生于 20 世纪 90 年代，主要用于传染 . com 文件。源码型病毒属于利用 C 语言、Java 语言编写的高级语言病毒，其攻击对象也是高级语言程序及其脚本。它在宿主程序编译之前进入源程序，在宿主程序编译之后与宿主程序融为一体，无法辨认。例如，Win32. Winux、Srevir 等病毒就属于这一范畴。操作系统型病毒主要工作在操作系统内部，通过攻击引导扇区和文件分配表实现攻击操作系统的目的，严重时可导致系统宕机。例如，Bouncing Ball 病毒、Stoned 病毒、Pakistan Brain 病毒等就是操作系统型病毒。

2）按传染原理分类

按传染原理分类就是按照病毒传染方式进行分类，大致可分为引导型、文件型、混合型病毒和蠕虫。引导型病毒寄居在计算机磁盘，通过攻击其引导扇区或主引导记录达到攻击的目的。它的执行先于计算机操作系统，在进行磁盘引导时便开始触发病毒，这样便会阻碍磁盘引导操作，待病毒代码开始运行之后，再进行磁盘引导。典型的病毒有 Brain、小球病毒等。文件型病毒通过感染 Windows 操作系统中的文件系统实现病毒的入侵。具体方式是感染扩展名为 . exe、. ovl、. com 的可执行类文件，当这些被感染文件开始执行时，病毒被触发，通常进行大量的自我复制，进而感染其他可执行文件，直至计算机内所有可执行文件都被感染。混合型病毒是引导型病毒和文件型病毒的综合体，感染方式也是两者的结合，它会同时入侵文件系统与磁盘引导区，因此算法比上述两者都要复杂。这种混合型病毒比单一型病毒有更强的传染性和更高的存活率，属于最难查杀的病毒种类。Tequila 病毒、Amoeba 病毒等都属于该范畴。蠕虫是通过间接自我复制的方式到处传播的一种病毒。它一般通过拦截世界各地的E－mail 文件传播自身的复制品。与病毒一样，蠕虫的破坏性是惊人的，传播速度更是极其迅速。例如，mellisa 病毒就属于这类病毒。

3）按破坏程度分类

按破坏程度分类，顾名思义就是按病毒对计算机系统的破坏性来划分的一种分类方法，

主要有良性病毒、恶性病毒之分。良性病毒不会对系统造成直接的破坏，也没有恶意代码的出现；恶性病毒则与之相反，通过恶意代码对系统造成直接的破坏。

11.1.3　计算机病毒检测技术

计算机病毒检测技术主要包括特征代码法、校验和法、行为监测法、感染实验法、长度检测法、软件模拟法、病毒签名检测法、人工智能方法等。以上病毒检测技术在进行病毒检测时具有不同的检测原理、检测范围，具体实现时也有不同的运算开销。其中，特征代码法、校验和法、行为监测法、感染实验法、长度检测法、软件模拟法、病毒签名检测法属于传统的病毒检测技术，而人工智能方法是一种新型的计算机病毒检测技术。

1. 传统的病毒检测技术

1）特征代码法

特征代码法早期被应用于多种计算机病毒检测工具中。特征代码法的方便之处是可快速地检测出网络中的已知病毒，将其应用于计算机网络进行病毒检测可获得最小的检测开销。特征代码法的实现步骤如下。

① 收集网络中已经出现过的病毒特征码，如果一种病毒同时感染了两种不同类型的文件，如 .com 文件和 .exe 文件，则在收集病毒特征码时要同时收集 .com 和 .exe 两种不同类型文件的病毒样本。

② 对于已经抽取的 .com 和 .exe 类型文件的病毒样本，还需要抽取这两种病毒样本中共有的病毒特征码，再把抽取到的相同代码加入病毒特征库。

③ 将所要检测的文件打开，启用文件的搜索功能，搜索文件中是否有与病毒特征库中的病毒特征码相匹配的代码，如果发现有与病毒特征码相匹配的程序代码，则可判断出被检测的文件已经被感染了哪种病毒。具备快速的病毒检测效率、较低的误报率且能够判断出是何种病毒等特点均是特征代码法的优势所在；然而其缺点也是不容忽视的，它无法识别已知病毒的变种及未知病毒，且由于要先抽取已知病毒的特征码，因此所需开销比较大，此外，因特征码的匹配需要很长时间，因此也会降低网络性能。

2）校验和法

对于系统中的正常文件，分别计算它们的校验和，并将计算所得的文件校验和写入该文件或将其写入一个统一的文件中进行保存，在需要启用该文件时，再次计算当前文件的校验和并将其与之前已保存的该文件的校验和进行比对，如果比较的结果不相同，则可断定该文件已经感染了病毒。通过以上描述的方法进行病毒检测即是校验和法，利用该方法既可检测出已知病毒，又能识别出未知病毒。利用校验和法进行病毒检测的优点是简单易用、可以识别出部分未知病毒，能够发现被检测文件发生的细小变化。相对于以上优点，该方法的缺点有：不可以确定病毒的名字及类型；误警率比较高，这主要是因为文件并不会只因为受到病毒攻击而发生变化，也有可能是因为部分软件版本的更新、口令的更改、一些运行参数的变化等情况；此外，校验和法无法检测出隐蔽型病毒，因为隐蔽型病毒进入系统内存后，会利用一定的方法清除感染病毒程序的病毒特征码，使得已经感染病毒的文件计算出与正常文件相同的校验和。

3）行为监测法

行为监测法即通过计算机病毒所独有的行为特征来监视病毒的检测方法。计算机反病毒研究人员通过对病毒程序长期的研究分析后发现，计算机病毒具备一些相同的行为特征，这些行为特征与正常的程序所表现出的特征不相同，且正常的程序文件中不会发现这些病毒行为。因此，可通过监测程序运行时所具有的一些行为特征来判断该程序是否已经感染病毒。如果监测出该程序出现与病毒行为类似的异常行为则说明已经感染病毒。能够识别网络中大部分的未知病毒是行为监测法的主要优点之一。但由于检测工具相当敏感，因此该方法容易产生误报现象；此外，该方法实现起来存在一定的困难，且不能判断出病毒的类型。

4）感染实验法

感染实验法主要的检测机理是利用病毒的感染特性。感染特性是计算机病毒的一个基本特性，所有病毒都会通过感染进行传播。该方法使用起来相对简单且比较实用。当网络中出现异常行为，且采用较新的检测方法也无法识别出该病毒时便可以采用感染实验法。首先运行一次可疑程序，再运行一些已经明确知道没有感染病毒的正常文件，然后对所运行的正常程序的长度及校验和进行观察，只要发现运行后的正常程序长度出现增加现象或校验和与之前计算的不相同，则可判断出该程序已经感染病毒。

5）长度检测法

感染特性是计算机病毒的基本特征，正常文件受到病毒感染后会明显增长，一般大概会增加几百字节。在目前的计算机系统中，文件的长度增加这一变化不太容易引起用户的关注，然而正常文件的长度在没有任何用户操作的情况下变长却是文件已经感染病毒的主要症状。长度检测法首先记录下正常文件的长度，并在文件被执行的过程中不断观察其长度是否发生变化，如果文件长度莫名其妙地增长则可判断该文件已经感染病毒。如果已知不同病毒使文件增长的具体长度，就可通过感染病毒文件增加的长度来大致判断该程序感染了何种病毒。这种根据文件增长的字节数来判断文件是否感染病毒的方法，在很多场合是有效的。但目前还没有哪种方法可检测出所有病毒。长度检测法仅通过可疑文件长度变化进行检测是不充分的。

6）软件模拟法

软件模拟法也称仿真扫描法，主要用来对付多态性病毒，因为多态性病毒每次都通过变化其病毒密码进行感染，每次感染时，都会以不同的随机数将自身加密到每个感染文件中，使中毒的正常文件出现不同的特点，同时病毒码也发生变化。因此，传统的病毒检测法无法发现这种病毒。为了实现对网络中出现的多态性病毒的检测，计算机病毒检测技术的研究人员专门研究出了软件模拟法。该方法主要是采用虚拟机技术实现的，用一个虚拟机仿真CPU、内存处理系统等一些系统组件，通过在虚拟机上模拟病毒代码的执行过程，并将其解密，提取特征字符串作为此病毒的特征码，从而掌握新病毒的结构及类型，并为制定杀毒措施提供条件，当再次发现这种病毒时，便可通过特征代码法加以匹配识别这种病毒。

7）病毒签名检测法

病毒签名是指宿主程序已经被感染病毒，它是给已经感染了病毒的程序文件所做的一种标记。类型不相同的病毒在对宿主程序进行感染时，会在宿主程序的不同地方嵌入不同感染

标记。感染标记主要由一些字符串或者数字串组成，如 FLU、1321、MSDOS、1753 等。不同程序在进行病毒签名时嵌入的内容、位置均不相同。通过对病毒样本文件进行分析，了解病毒签名嵌入的内容及位置以后，根据嵌入的特定位置搜索可疑程序中的病毒签名。如果在该位置搜索到了一种病毒签名，则可判断出可疑程序中存在病毒，且可根据签名嵌入位置判断出是哪种病毒。该方法即为病毒签名法。该方法的特点是首先必须知道该种病毒签名的内容和位置，要掌握各种病毒的签名，必须对病毒进行分析，且容易出现虚假报警。

2. 新型病毒检测技术

随着计算机病毒的发展，特别是变形病毒和网络病毒肆虐，传统的病毒检测技术及一些高级病毒检测技术在检测范围等诸多方面的局限性也慢慢体现出来，而人工智能的研究启发人们将人工智能的方法应用于计算机病毒检测系统，研究者主要希望通过人工智能技术自动地抽取计算机病毒特征，并自动对病毒进行检测，因此基于人工智能的新型病毒检测方法应运而生。目前人工智能领域用于病毒检测的技术主要有：利用人工免疫的思想进行病毒检测、基于神经网络的病毒检测技术、通过数据挖掘方法检测恶意代码、应用机器学习算法对计算机病毒进行检测。人工智能特别是人工免疫系统的研究引发人们思考，面对网络中层出不穷的变形病毒，研究计算机病毒检测技术的核心是如何快速地识别出网络中的合法程序（即自我信息）及非法程序（即非我信息）。而生物免疫系统能精确识别出生物体内的“自体”及“非自体”信息的特点启发了人们，将人工免疫理论应用到计算机病毒检测领域能为病毒检测及防止黑客攻击提供一个新途径。

11.2　被动防御——防火墙技术

11.2.1　防火墙的概念

防火墙是位于两个（或多个）网络间实施网间访问控制的一组组件的集合，它能增强机构内部网络的安全性。防火墙系统决定了哪些内部服务可以被外界访问，哪些外部服务可以被内部人员访问。

防火墙必须满足以下三个条件：

- 所有进出被保护网络的通信必须通过防火墙；
- 所有通过防火墙的通信必须经过安全策略的过滤或防火墙的授权；
- 防火墙自身应对渗透免疫。

11.2.2　防火墙的分类

根据防火墙的组成、实现技术和应用环境等方面的不同，可以对防火墙进行分类。根据防火墙组成组件的不同，可以将防火墙分为软件防火墙和硬件防火墙。软件防火墙以纯软件的方式表现，安装在边界计算机或服务器上就可以实现防火墙的各种功能。硬件防火墙以专用硬件设备形式出现，一般情况下硬件防火墙都以软件和硬件相结合的方式实现。它根据设计需求选用合适性能的硬件，然后再按照设计安装上选定的操作系统和软件防火墙系统，有时软件防火墙系统会和操作系统集成在一起。完全通过硬件实现的防火墙系统是防火墙技术

的一个发展方向，国外已有不错的产品。软件防火墙的特点是成本低、性能也较低，一般适合于规模较小或对外带宽较窄的网络系统；硬件防火墙的特点正好相反，完全通过硬件实现的硬件防火墙则能提供更高的性能指标。硬件防火墙是防火墙产品的主流。

根据防火墙的实现平台，防火墙可分为基于 Windows 平台的 Windows 防火墙和基于 Linux 平台的 Linux 防火墙等。软硬结合的硬件防火墙一般在内置的 Linux 平台上实现，而软件防火墙一般需要支持的平台比较多。根据平台操作系统自身的复杂性和代码开放程度不同，防火墙研发的难度相差较大，Linux 防火墙应用很广，其实现却相对容易，Windows 防火墙则正好相反。

防火墙分类如图 11-2 所示。根据防护对象的不同，防火墙可以分为主机防火墙和网络防火墙。传统防火墙都是网络防火墙，主机防火墙（也称个人防火墙或 PC 防火墙）的设计还是最近几年的事情。

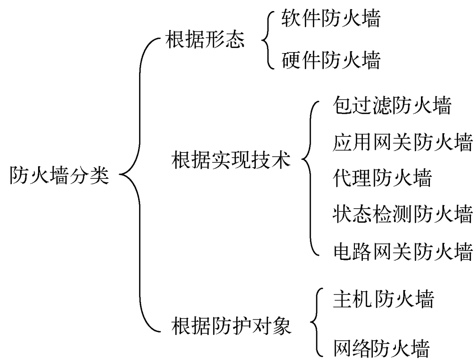


图 11-2　防火墙分类

根据防火墙自身网络性能和被保护网络系统的网络性能，防火墙又分为百兆防火墙和千兆防火墙。百兆防火墙能够提供百兆比特带宽的网络接口，适合于出口带宽在百兆以内的网络系统。千兆防火墙至少提供一个一千兆比特带宽的网络接口，适合于出口带宽高于百兆网络系统的安全防护。对于国家主干网络之间或大型内网保护，万兆防火墙也是必要的。不同硬件厂商提供的面向千兆防火墙的硬件平台各有优缺点，或者性能稳定，或者兼容性好，而且都不能很好地支持高速加密服务。千兆防火墙是目前防火墙技术发展的焦点，应该说目前千兆防火墙应用技术并不是很成熟。

根据防火墙功能或技术特点，防火墙又包括主机防火墙、病毒防火墙和智能防火墙等。防火墙体系结构和实现技术是防火墙最常用的分类依据。

根据防火墙自身的体系结构，可以将防火墙分为以下种类：包过滤防火墙、应用层代理、电路级网关、地址翻译防火墙和状态检测防火墙等。

11.2.3　防火墙的配置

防火墙是网络安全的一个重要防护措施，用于对网络和系统的保护。监控通过防火墙的数据，根据管理员的要求，允许和禁止特定数据包的通过，并对所有事件进行监控和记录。如何配置一个高效的防火墙是一个值得关注的问题。

目前主流的防火墙配置方案有三种：双宿主机网关、屏蔽主机网关、屏蔽子网。

1. 双宿主机网关

这种配置是用一台装有两个网络适配器的双宿主机做防火墙。双宿主机用两个网络适配器分别连接两个网络，又称堡垒主机。堡垒主机上运行着防火墙软件（通常是代理服务器），可以转发应用程序，提供服务等（见图 11-3）。

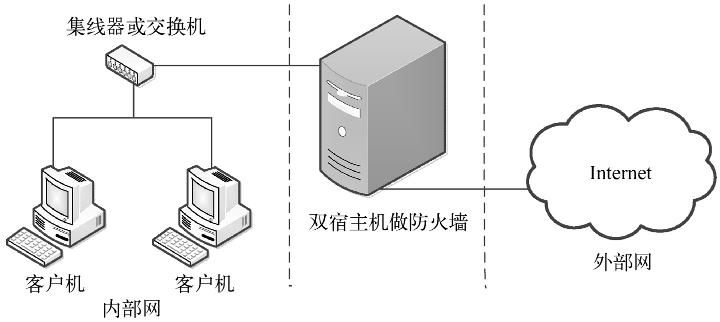


图 11-3　双宿主机网关

双宿主机是唯一隔开内部网和外部 Internet 之间的屏障，如果入侵者得到了双宿主机的访问权，内部网络就会被入侵，因此为了保证内部网的安全，双宿主机应该具有强大的身份认证系统，才可以阻挡来自外部不可信网络的非法登录。

为了防止防火墙被入侵，在系统中应尽量减少防火墙上用户的账户数目。使用双宿主机应该注意的是，首先要禁止网络层的路由功能。由于双宿主机是外部用户访问内部网络系统的中间转接点，它必须支持很多用户的访问，因此双宿主机的性能非常重要。

2. 屏蔽主机网关

屏蔽主机网关易于实现，安全性好，应用广泛。它又分为单宿堡垒主机和双宿堡垒主机两种类型。

单宿堡垒主机类型，由一个包过滤路由器连接外部网络，同时一个堡垒主机安装在内部网络上。堡垒主机只有一个网卡，与内部网络连接（见图 11-4）。通常在路由器上设立过滤规则，并使这个单宿堡垒主机成为从 Internet 唯一可以访问的主机，确保了内部网络不受未被授权的外部用户的攻击。而 Intranet 内部的客户机，可以受控制地通过屏蔽主机和路由器访问 Internet。

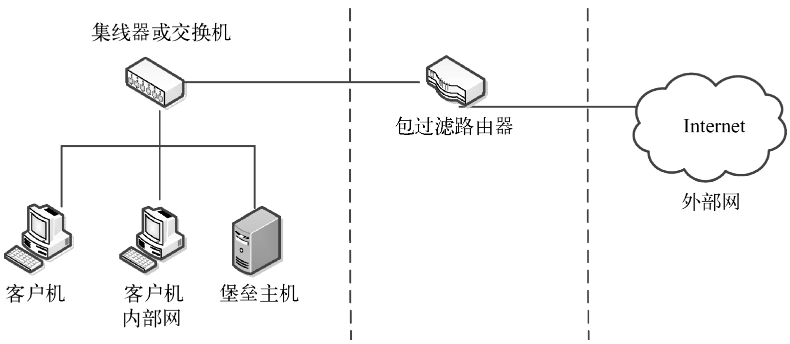


图 11-4　屏蔽主机网关（单宿堡垒主机）

双宿堡垒主机型与单宿堡垒主机型的区别是，堡垒主机有两块网卡，一块连接内部网络，一块连接包过滤路由器（见图 11-5）。双宿堡垒主机在应用层提供代理服务，与单宿型相比更加安全。

在采用屏蔽主机网关的情况下，包过滤路由器是否正确配置是这种防火墙安全与否的关

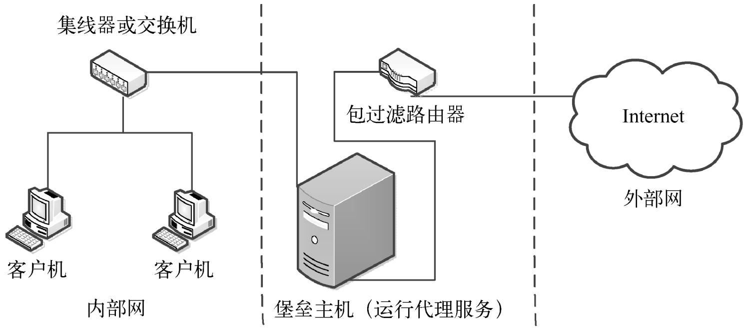


图 11-5　屏蔽主机网关（双宿堡垒主机）

键，过滤路由器的路由表应当受到严格的保护，否则如果路由表遭到破坏，数据包就不会被路由到堡垒主机上，使堡垒主机被越过。

3. 屏蔽子网

屏蔽子网是在 Intranet 和 Internet 之间建立一个被隔离的子网，用两个包过滤路由器将这一子网分别与 Intranet 和 Internet 分开。两个包过滤路由器放在子网的两端，在子网内构成一个“缓冲地带”（又称“非军事区”，如图 11-6 所示），两个路由器中的一个控制 Intranet 数据流，另一个控制 Internet 数据流，Intranet 和 Internet 均可访问屏蔽子网，但禁止它们穿过屏蔽子网通信。可根据需要在屏蔽子网中安装堡垒主机，为内部网络和外部网络的互相访问提供代理服务，但是来自两个网络的访问都必须通过两个包过滤路由器的检查。对于向 Internet 公开的服务器，像 WWW、FTP、Mail 等 Internet 服务器也可安装在屏蔽子网内，这样无论是外部用户，还是内部用户都可访问。

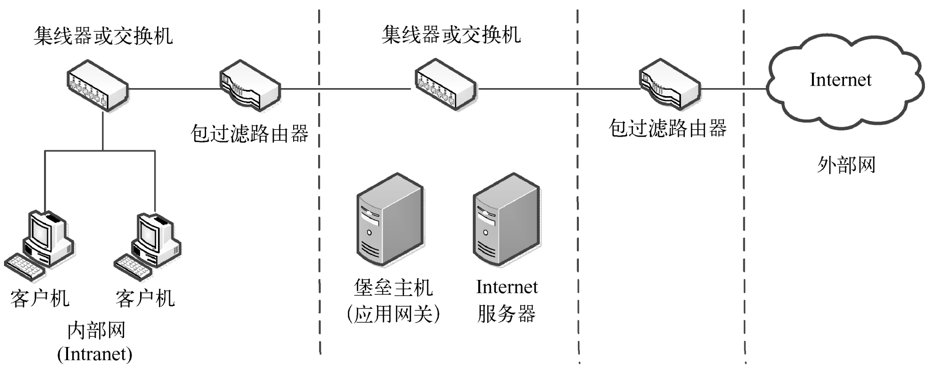


图 11-6　屏蔽子网防火墙

屏蔽子网防火墙的安全性能高，入侵者需要突破 3 个不同的设备才能入侵内部网络；只对外通告 DMZ 区的网络，保证内部网络不可见；内部网络用户通过堡垒主机或代理服务器访问外部网络。但其需要的设备较多，造价较高。

11.3　主动防御——入侵检测技术

从计算机系统和网络安全的目标来看，入侵是一种试图破坏网络或系统资源可用性、完

整性和保密性的行为，或是一种违背系统安全策略的事件。从入侵策略的实施角度进行分析，入侵又可分为内部的冒充合法用户和外部的企图进入和成功闯入等多个方面。简单来说，入侵检测（Intrusion Detection，ID）就是主动检测入侵事件的安全策略技术。根据美国安全通信委员会（NSTAC）在 1997 年所作的关于“入侵检测”的定义，可以将入侵检测理解为检测并识别处于不同实施阶段的入侵行为的过程，发现处于企图/试图、正在进行和已经完成的入侵行为。入侵检测系统则是能够实现入侵检测功能的软硬件结合的系统，该系统从网络或计算机系统内部的若干关键节点获取信息（包括网络行为、安全日志、审计数据等信息），分析数据从而发现是否有违反安全策略的行为和攻击迹象，并做出相应响应。作为一种主动安全防御技术，入侵检测能够检测内外部攻击和误操作，在系统受到危害之前拦截入侵以提供实时保护。

11.3.1　IDS 的标准结构

根据目前国际通行的方法，可以将入侵检测系统（Intrusion Detection System，IDS）的内部结构分为如下几个部分：事件产生器（EventGenerators）、事件分析器（EventAnalyzers）、响应单元（ResponseUnits）和事件数据库（EventDatabases）。

事件产生器又称传感器，是入侵检测的第一步，它的目的是从整个网络环境中采集数据，并提供给系统的其他部分。采集内容可能包括：系统日志、应用程序日志、系统调用、网络数据、用户行为和其他 IDS 的信息。从技术上来说，传感器实际就是一个嗅探器（Sniffer）。

事件分析器分析得到的数据，并产生分析结果。它是整个 IDS 的核心，效率高低直接决定整个 IDS 的性能。

响应单元则是对分析结果做出反应的功能单元，功能包括：

① 报警和事件报告；

② 终止进程，强制用户退出；

③ 切断网络连接修改防火墙设置；

④ 灾难评估，自动恢复；

⑤ 查找定位攻击者。

事件数据库是存放各种中间和最终数据的地方的统称。它可以是复杂的数据库，也可以是简单的文本文件。通常，规则库、行为模式库等也归于此。

此外，一套完整的 IDS 还应包括管理器，它负责定位、控制等常规的管理功能，包括管理员控制台和日志输出模块。管理员控制台可以采用命令行或 Web 方式。

11.3.2　IDS 的分类

IDS 系统的分类标准很多。从数据分析手段看，入侵检测分为两类：误用（Misuse）检测和异常（Anomaly）检测。从数据来源看，入侵检测分为两类：基于主机的入侵检测和基于网络的入侵检测。关于入侵检测技术的分类如图 11-7 所示。

1. 误用检测

误用检测建立在对过去各种已知网络入侵方法的系统缺陷知识的积累之上，它需要首先

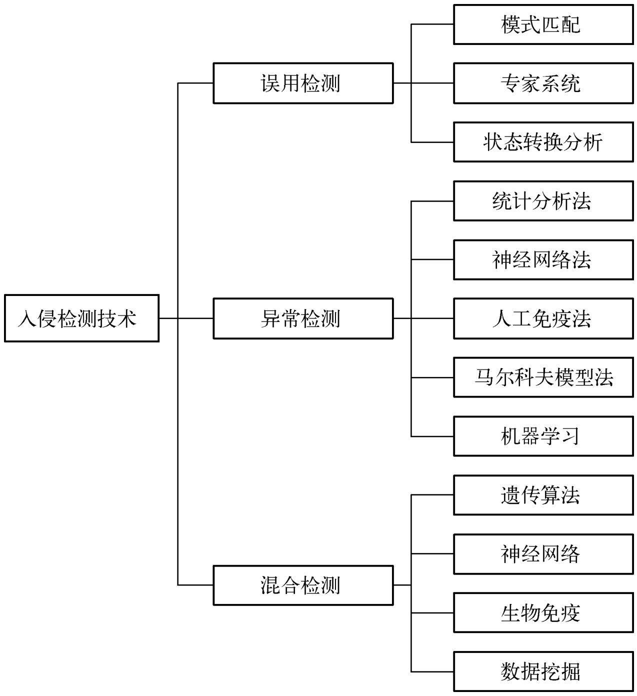


图 11-7　入侵检测技术分类

建立一个包含上述已知信息的数据库，然后在收集到的网络活动信息中寻找与数据库项目匹配相关的蛛丝马迹。此技术的优点在于其较高的检测率和较低的误警率，缺点是不能发现未知的攻击。误用检测的实现方法如下。

① 模式匹配：模式匹配主要是用一定的模式描述来提取攻击的主要特征，其基本任务就是把存放在入侵检测规则集中的已知入侵模式与系统正在检测的网络包或重构的 TCP 流中的文本进行匹配，如果匹配成功，则可以断定发生了入侵。这个过程是不断循环进行的，它具有较高的检测率和较低的误警率，其检测规则必须不断地更新。

② 专家系统：用专家系统对入侵进行检测，经常是针对有特征的入侵行为。它将有关入侵的知识转化为 if－then 结构，if 部分为入侵特征，then 部分为系统防范措施。所谓的规则即知识，专家系统的建立依赖于知识库的完备性，知识库的完备性又取决于审计记录的完备性与实时性。专家系统的难点就在于实现专家系统知识库的完备性。

③ 状态转换分析：它将入侵检测表示成一系列被监控的系统状态迁移，攻击模式的状态对应于系统状态，并具有迁移到另外状态的条件判断。通过弧将连续的状态连接起来以表示状态改变所需的事件，允许事件类型被植入模型并且无须与审计记录一一对应。

2. 异常检测

这种方法是建立在如下假设基础之上，即任何一种入侵行为都能由于其偏离正常或所期望的系统和用户的活动规律而被检测出来。此技术的优点在于理论上能够发现任何攻击行为，缺点是误警率较高。

① 统计分析法：统计分析是一种广泛被使用的入侵检测技术。它通过在一段时间内收

集与合法用户行为相关的数据来定义正常的阈值（Threshold），如果当前的行为偏离了正常行为的域值，就表示有入侵发生。对于用户所生成的每一个审计记录，系统经计算生成一个单独的检测统计值 T^2，用来综合表明最近用户行为的异常程度。较大的 T^2 值将指示有异常行为的发生，而接近于零的 T^2 值则指示正常的行为。统计值 T^2 本身是一个多个测量值异常度的综合评价指标。假设有 n 个测量值表示为 S_i,$(1\leqslant i\leqslant n)$，则 $T^2=a_1S_1^2+a_2S_2^2+\cdots+a_nS_n^2$，其中 $a_i(1\leqslant i\leqslant n)$ 表示第 i 个测量值的权重。

其优点是检测率较高，因为可以使用不同类型的审计数据，缺点是它对一些行为不很敏感，并且可检测到的入侵类型也受到限制。

② 基于神经网络的异常检测方法：神经网络模仿生物神经系统，通过接收外部输入的刺激，不断获得并积累知识，进而具有一定的判断预测能力。这种方法通过对一个特定用户先前命令序列的分析推测出下面要执行的命令。它包括三个阶段：一是在一定的时期从每一个用户的审计日志中来收集训练数据，这样就会形成一个向量，以表明每个用户多长时间执行一条命令；二是在命令分布向量的基础上训练神经网络以标识用户；三是在运行过程中，使用神经网络识别新一天的向量，如果网络表示和实际用户有很大的不同或没有一个明确的建议，就表示有异常行为发生。

基于神经网络的异常入侵检测系统具有学习的能力，它可以紧密地模仿用户的行为并且根据最近的变化进行调整。它的另外一个特性就是允许模糊数据或噪声数据。此外，与统计理论相比，神经网络更好地表达了变量之间的非线性关系。其缺点是需要的计算负载较重，并且很难解释输入和输出之间的关系。

③ 基于人工免疫的异常检测：将被检测网络中正常活动视为自我，异常活动视为非我，其目的就是区分正常或异常的网络活动。人工免疫模型的工作流程分为三个阶段，即生成规则基因库、筛选检测规则集和复制高效检测规则集。该入侵模型可以分成两个检测层次，一个是系统级检测层次，另一个是网络级检测层次。在系统级检测层中主要监控主机的各种操作行为。用户的删除、修改、格式化等操作都要接受该层的分析和识别。而网络级检测层主要负责对网络上传输的数据的监控，包括网络数据包的识别和检测、地址的过滤等。

人工免疫系统归根结底是进行“自我”和“非我”的识别。而在该入侵检测模型中，把与所需检测的机器相连的网络间正常的 TCP/IP 连接集合和该机器系统内合法的操作行为定义为“自我”，采用可以描述 TCP/IP 连接的特征信息来表示。例如，源 IP 地址、目的 IP 地址、服务端 VI、协议类型、包的数量、字节数、特定错误和在短时间内网络的特定服务，以及描述系统合法操作的集合等。而把反常的 TCP/IP 连接集合和非法的系统操作集合定义为“非我”，而这些特征信息在具体表现形式上都可以通过某种规则映射为唯一表征该信息的长度为 1 的二进制字符串。该方法成功地将人工免疫理论应用到入侵检测中，但目前还只处于研究阶段。

④ 基于隐马尔科夫模型（HMM）的入侵检测方法：一个系统调用既可以是完全正常的，也可以是危险的。例如，被缓冲区溢出攻击的程序，它所产生的系统调用事件和正常情况下产生的有着明显的不同。因此，可以通过构建正常情况下的系统调用事件模型，然后观察是否与此模型有明显的偏离，以此来有效地检测入侵的产生。

隐马尔科夫模型是对观察到的符号序列构造模型的一种非常好的工具，它在构造系统调用事件模型上有着比其他方法更好的性能，但它在构造正常行为模型时需要较长的时间。解决办法是提高计算机系统的性能，或者减少观察的数据。

⑤ 机器学习：该方法通过对新序列（如离散数据流和无序的记录）的相似度的计算，将原始数据转化为可度量的空间，然后应用 IBL（Instance Based Learning）学习技术和一种新的基于序列的分类方法，发现异常事件，从而检测入侵行为。这种方法检测速率高，且误报率较低。然而，这种方法用于用户动态行为变化及单独异常检测时还有待改善。

3. 混合检测

使用单一的方法进行入侵检测受到一定的局限，要么不能检测未知入侵，要么检测率不高，达不到有效检测的目标。因此，使用多种检测方法来检测入侵受到研究人员的关注，目前已提出多种混合检测方法。

① 遗传算法：这种方法寻找已知攻击向量（这种向量的每个元素表示一个特定的攻击），能最好地匹配观察到的事件流。基于相关攻击的危害度和不匹配部分的二次惩罚函数来估计假设向量。在每一代中，经过交叉和变异得到当前最好假设向量，以使误报率和漏报率都趋于零。这种方法具有较好的性能，但它不能辨别攻击匹配的原因，而且也不能表示同时或组合攻击。

② 神经网络：这种方法利用神经网络技术来进行入侵检测。它用于检测未知攻击，使用神经网络对输入向量（来自于审计日志或正常的网络访问行为，经数据信息预处理模块的处理）进行处理，从中提取用户正常行为的模式特征，并以此创建用户的行为特征轮廓。因此，这种方法对用户行为动态性要求入侵检测系统具有学习和自适应功能，能够根据实际检测到的信息有效地加以处理并做出入侵可能性的判断。

③ 生物免疫：生物免疫系统是为了保护个体（自己）不受故意生物（“异己”）的侵害，而入侵检测则为保护一台或一组计算机不受入侵，因此生物免疫系统的原理、算法和体系结构均可用于入侵检测。基于免疫学的入侵检测系统模仿生物免疫系统，实时监控和处理主机的审计数据，提取感兴趣的行为数据，建立行为特征模式，并与用户的行为模式匹配，以此来发现是否发生入侵。这种模型的主要思想是区分“自我”和“非我”。“自我”是正常行为，“非我”指异常行为。根据生物免疫的基本思想，将正常网络访问当成正常行为，将异常访问当成异已行为，区别自己和异己从而判别出入侵行为，当一个行为模式和一个“异己”模式很相近或相同时可判别发生入侵行为。

④ 数据挖掘：用数据挖掘程序处理搜集到的审计数据，为各种入侵行为和正常操作建立精确的行为模式，这是一个自动的过程，不需要人工分析和编码入侵模式。数据挖掘方法的关键点在于算法（分类算法、关联算法和序列分析）的选取和建立及一个正确的体系结构。

入侵检测作为一种积极主动的安全防护技术，提供了对内部攻击、外部攻击和误操作的实时保护，在网络系统受到危害之前拦截和响应入侵。误用检测具有较高的检测率，但不能检测新型入侵，异常检测能检测新型入侵，但误报率较高，发展的趋势是采用计算智能等检测方法，以实现自动化、分布式入侵检测，提高检测率，降低误报率。

11.4 主动防御——网络态势感知技术

Bass 于 1999 年首次提出网络态势感知（Cyberspace Situational Awareness，CSA）的概念，并且指出“基于融合的网络态势感知”必将成为网络管理的发展方向。所谓网络态势是指由各种网络设备运行状况、网络行为及用户行为等因素构成的整个网络的当前状态和变

化趋势。值得注意的是，态势强调环境、动态性及实体间的关系，是一种状态、一种趋势、一个整体和宏观的概念，任何单一的情况或状态都不能称为态势。CSA 是指在大规模网络环境中，对能够引起网络态势发生变化的要素进行获取、理解、评估、显示及对未来发展趋势进行预测，因此这里将其看成一种主动的防御技术。

态势感知的概念源于军事需求，是军事决策制定过程的重要环节。CSA 的目标是将态势感知的成熟理论和技术应用于网络管理，在急剧动态变化的复杂环境中，高效组织各种信息。将已有的表示网络局部特征的指标综合化，使其能够表示网络的宏观、整体状态，从而加强管理员对网络的理解能力，为高层指挥人员提供决策支持。与传统的网络管理系统相比，CSA 具有以下特点：

① 运用数据融合技术，综合考虑影响网络态势的多种因素，提供全面而宏观的网络状态视图，加强对网络的理解与控制，以减轻网络管理员的负担；

② 成为集成单元网管的平台，改变当前各单元网管独立工作的局面，实现信息共享，海量多样的信息带来丰富而准确的分析结果；

③ 为风险评估、决策制定提供支持。

一旦完成态势感知，决策几乎可以根据态势自动生成。参考龚正虎等人的研究论文，下面给出态势感知的后续内容。

11.4.1　网络态势感知研究框架

CSA 作为数据融合的一部分，并不是孤立存在的，向下从 Level 1 融合获取各类网管数据，向上为 Level 3 融合提供态势信息，用于威胁分析和决策支持，而且与其他融合层次关系紧密。层与层之间不仅数据通信频繁，而且方法相通，没有明确的界限，作为一个整体而存在。因此，CSA 研究包括多方面内容，其总体研究框架如图 11-8 所示。CSA 研究框架概

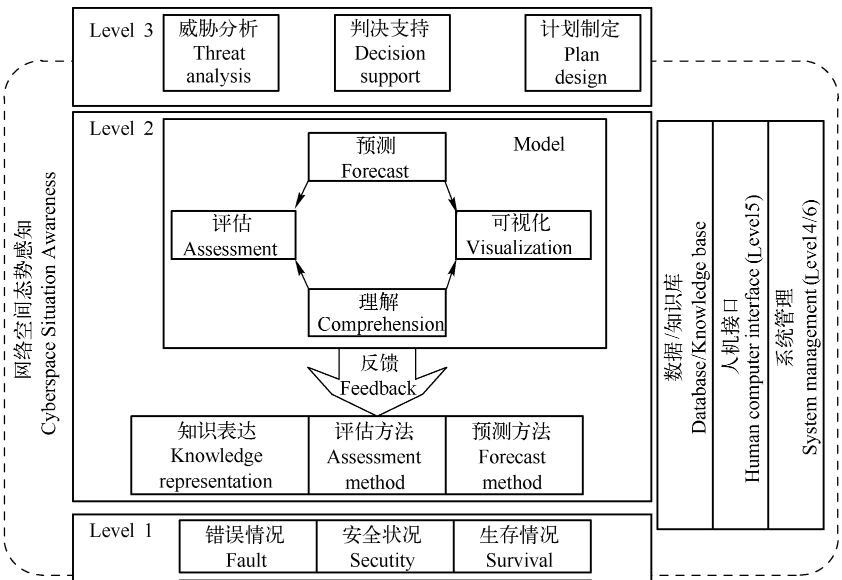


图 11-8　网络态势感知研究框架

括 CSA 研究内容，体现 closing - the - loop 理念，突出动态循环、不断细化的本质，强调反馈的重要作用。由图 11-8 可知，CSA 研究内容广泛，包括自身功能细化、关键技术的理论方法、与其他融合层次的通信与交互。目前的研究主要集中在 3 个方面：CSA 模型、知识表示和评估方法。其中：模型是研究的重点，相对比较成熟、统一；评估方法是 CSA 研究的核心，主要研究现有理论在态势评估中的运用；而有关知识表示的研究相对较少。根据不同的应用领域，网络态势又可分为安全态势、拓扑态势、传输态势、生存性态势等，从 Bass 开始，多数研究围绕安全态势展开。

11.4.2 网络态势感知模型

态势感知以数据融合为中心，其模型的建立也以数据融合模型为基础。目前已经提出了几十种数据融合模型：智能循环、JDL 模型、Boyd 控制循环、Endsley 模型、瀑布模型、Dasarathy 模型、Omnibus 模型、扩展 OODA（Observe，Orient，Decide and Act）模型，此外还有知觉推理模型，依赖联想存储和关联数据库，模仿人类思维方式。

其中最具影响力的当属 JDL 数据融合模型。JDL 模型将融合分为 4 个层次：目标细化、态势细化、风险细化、过程细化，其中：态势感知作为较高层次的 Level 2 融合，向下从 Level 1 融合接收网络元素的监测数据，作为态势感知的信息来源；向上为 Level 3 融合提供态势信息，用于威胁分析和决策支持。Blasch 等学者对 JDL 模型加以发展，在 4 层融合的基础上提出了第 5 层——用户细化。Level 5 融合强调用户的作用，需要用户的知识和推理；同时，为低层提供反馈，优化融合过程。目前，JDL 模型进一步发展为 DFIG（Data Fusion Information Group）模型，DFIG 提出了第 6 层——任务管理，区分信息融合和管理功能。

Boyd 控制循环模型（OODA），描述了目的与活动的感知过程，并将感知循环过程分为观察、判断、决策、行动这 4 个阶段。其中：观察从物理域跨越到信息域，判断和决策属于认知域；而行动从信息域回到物理域，完成循环。前 3 个阶段类似于 JDL 模型，行动阶段通过考虑决策在真实世界中的影响来闭合循环。扩展 OODA 模型提供了处理多并发、潜在交互数据融合的机制。

Endsley 模型对态势感知阶段做了进一步细化，分为态势觉察、态势理解、态势预测 3 个层次。目前，Endsley 模型日渐受到关注。

在通用态势感知参考模型的基础上，结合网络应用具体问题提出了 Cyber SA 模型。Cyber SA 模型包含日志、配置、任务、攻击、入侵尝试等网络元素，体现出网络的特点。尽管态势感知模型数目繁多，但是从中可以看到以下几点共性。

① 模型的重点放在态势感知的功能上。除 Dasarathy 模型关注任务之外，其他模型都是对数据融合，进而对态势感知的功能进行划分。不同模型的组成部分尽管名称不同，但很多功能是一致的。

② 循环是态势感知的本质。模型中的每个组成部分没有明确的顺序，而是重复迭代进行。

③ 强调反馈的作用。多数模型最终形成闭环系统。

11.4.3　网络态势知识表示

知识表示解决两个问题。第一个问题是对不确定信息的表示。信息表示关注的焦点是不确定性。信息不确定性的表示绝非态势感知所特有，相关的研究有很多，而且作为基础研究，提出了从表示到融合完整通用的解决方案，其中常见的不确定性理论有概率论、模糊集、可能性理论和证据理论等，已经在态势感知领域得到了应用。而且，不确定性本身的形式很多，在不同的领域中其分类和定义也不一样。从语义上讲，基本上与 Klir 和 Yuan 的论文“Fuzzy Sets and Fuzzy Logic”的分类相差无几，即把不确定性分为模糊性（fuzzy）和多义性（ambiguity），而多义性又可分为非特异性（nonspecificity）和冲突（conflict）。对应这些类型的不确定性，不同的不确定性理论所能处理的不确定性的种类也不一样。模糊集是处理模糊性的理论，概率论只涉及事件之间的冲突；可能性理论表示出事件的非特异性，而证据理论描述了非特异性和冲突。其中，使用最多的当数 Zadeh 教授提出的模糊集，是处理人类认知不确定性和不精确性问题的数学方法，可以对不确定的语义信息进行处理，是描述人脑思维处理模糊信息的有力工具；它不仅可以独立使用构建表示和推理系统，还可以和其他技术相结合作为通用的信息表示技术。模糊集在态势感知领域得到认可并广泛使用。此外，灰色理论是一种利用灰色信息的理论方法，采用白化函数对各种灰色信息进行白化，从而得到灰色信息的量化估计。

另一个问题是对复杂系统的表示，这也是 CSA 的挑战之一。为了改变当前各单元网管独立工作的状况，建立真正意义上的 CSA，在很大程度上依赖统一的系统表示方法。能够对网络做出准确、全面、详尽的描述，是进行态势感知的前提。另一方面，网络作为复杂巨系统，面对纷繁丰富的内容和错综复杂的关系，对其进行描述的能力很弱，而相关的研究比较少。本体论是其中的重要方法。本体源于哲学概念，被引入计算机科学领域后，特指对共享概念模型所做的明确而规范的形式化说明。Husserl 于 1985 年提出了形式化本体论（formal ontology）。本体强调领域中的本质概念，同时也强调这些本质概念之间的关联，能够将该领域中的各种概念及概念之间的关系显式化、形式化地表达出来，从而表达出概念中包含的语义，增强对复杂系统的表示能力。

11.4.4　评估方法分类

作为 CSA 研究的核心，评估方法多种多样，这些方法大致可以分为 3 类，而对于具体的评估方法，这里不做介绍。

1. 基于数学模型的方法

基于数学模型的方法（Mathematics Model，MM）就是综合考虑各项态势因子，构造评定函数，建立态势因子集合 R 到态势空间 θ 的映射关系 $\theta=f(r_1,r_2,r_3,\cdots,r_n)$，其中 $r_i\in R(1\leqslant i\leqslant n)$ 为态势因子。态势包含众多相互冲突、不可公度、不确定的复杂态势因子，这些因子具有层次结构，可以逐层划分、细化。传统、通用的多目标决策理论和效用理论的有关方法，如最大隶属原则、距离偏差法、打分法、多属性效用函数等，都可以用于态势评估。最常用的当

属权重分析法及公式法和集对分析。

MM 方法能建立明晰的数学表达式，模型易于理解，而且能够建立连续的态势空间，给出一种有利或不利的判断性结果，便于态势的优劣对比。但是评定函数的构造、参数的选择没有统一科学的方法，一般依赖领域知识和专家经验，不可避免地带有主观意见，缺少科学客观的依据。此外，态势评估多数情况使用自然语言表述知识，而这种知识不容易被转化为易于机器处理的数学表达式。因此，建立面向自然语言条件陈述的数学模型也成为该方法的难点。

2. 基于知识推理的方法

基于知识推理的方法（Knowledge Reasoning，KR）的基本思路是在已知经验知识、先验概率的前提下，接收 Level 1 融合的输出，根据实时监测的数据信息，通过一定的关系逐级推理得到对当前态势的判断，并且可以对态势空间进行划分，给出分级或分类结果。KR 方法又可以分为基于产生式规则的逻辑推理、基于图模型的推理和基于证据理论的概率推理，代表性的方法有模糊推理、贝叶斯网络、马尔科夫过程、D－S 证据理论。KR 方法能够模拟人类思维方式，较之 MM 方法，将知识的运用融入推理的过程之中，具有一定的智能，类似于专家解决问题的过程。评估的结果建立离散的态势空间，能够确定态势优劣等级，或者能够指明态势的类型，一目了然，便于态势的理解和把握。该方法的难点在于如何获取建立模型所需的知识。凭借经验会带有强烈的主观性，如果通过机器学习，相关的研究还比较少，而且如何学习到“模拟人类思维方式”的知识更是难上加难。可见，该方法的优点反而成了最大的障碍。此外，该方法维护大量推理规则，空间开销和推理代价都很高。如何应对大规模的问题是另一个需要考虑的问题。

3. 基于模式识别的方法

基于模式识别的方法（Pattern Recognition，PR）分为建立模版和模式匹配两个阶段。第 1 阶段建立态势模版，在对态势空间进行划分的基础上，识别所有可能出现的态势状态。划分没有统一的标准，可以将态势分为不同类型，也可以对态势进行分级。第 2 阶段，通过计算实测数据与模版数据之间的关联，如果两组数据符合，或者关联系数达到预先规定的阈值，则认为匹配成功，从而确定态势状态。建立模版是 PR 方法的重点，关键在于选择分类方法。除了凭借专家经验、领域知识以外，机器学习也是划分的主要手段，可从训练样本或案例中获得有关分类的知识，代表性的方法有基于案例的推理、神经网络、灰关联分析、粗集理论、聚类分析。这些方法一般也用于模式匹配。

PR 方法引入机器学习机制，科学、客观，可以方便地从历史数据或案例中获得态势划分的知识。该方法计算量大，在非实时环境中有很好的效果，但是在实时环境中可能无法满足要求，某些研究采用启发式算法可提高效率。此外，由于分类知识是从历史数据中通过机器学习获得的，机器很难给出直观的解释，不利于理解。

11.5　主动防御——移动目标防御技术

一些研究指出，互联网安全的最大问题在于整体态势的易攻难守：

① 攻击者具有时间优势来组织攻击，即攻击者可长期对攻击目标（网络设备、通信链路、基础协议等）的固有脆弱性进行反复的漏洞分析和渗透测试，直至达到最终目标；

② 攻击者具有信息不对称优势来发射攻击，即攻击者只需要找到一个有效攻击点即可实施攻击，而防御者则需要对所有可能出现的攻击点和攻击方式加以防护；

③ 攻击者具有成本优势来推广攻击，即攻击者一旦成功实施一次攻击，便可以以较低成本轻易将攻击范围扩大。

究其原因，可归结为当前网络系统的确定性、静态性和同构性：网络的确定性和静态性使攻击者具备时间优势和信息不对称优势，同构性则使攻击者具备成本优势，从而导致防御者始终处于被动的劣势地位。而这种被动劣势无法依靠现有防御方法来弥补：

① 由于人的认知有限性，常用的源代码检查机制难以保证能排除所有的漏洞；

② 补丁下发通常明显滞后于攻击者对安全漏洞的利用，这一时间差为网络攻击提供了生存空间；

③ 识别攻击代码与感染症状机制需要判断出攻击的签名或预先对恶意攻击进行定义，但攻击者速度快、灵敏度高且会使用简单的多态机制来持续改变攻击的签名，从而使防御方所采用的基于签名的方法在很大程度上都变得无效。

为改变网络空间中这种“易攻难守”的局面，美国政府开始重视发展“改变游戏规则”的革命性技术以实现积极主动的网络防御，并相继出台了一系列重要纲领性文件，均着重强调针对当前网络空间所面临的实际和潜在威胁而发展“改变游戏规则”革命性技术的必要性与紧迫性。移动目标防御（Moving Target Defense，MTD）就是美国针对防御者当前所处劣势地位而提出的一个“改变游戏规则”的网络安全发展方向，期望通过实施持续、动态的变化迷惑攻击者，以增加其攻击成本和复杂度，降低其攻击成功率。值得注意的是，移动目标防御不是某一种具体的防御方法，而是一种设计指导思想。这一思想可应用到被保护系统的某一属性上，衍生出多种具体的防御机制。在国内，邬江兴院士则提出了拟态安全防御思想，期望通过在主动和被动触发条件下动态、伪随机地选择执行各种硬件变体及相应的软件变体，使得内外部攻击者观察到的硬件执行环境和软件工作状况非常不确定，无法或很难构建起基于漏洞（bug）或后门的攻击链，以达成降低系统安全风险的目的。

11.5.1　移动目标、移动目标防御及拟态安全防御

对于移动目标，当前并没有一个权威的统一定义，仅在美国白宫国防安全委员会的进展报告中提到移动目标是可在多个维度上移动以降低攻击者优势并增加弹性的系统。对于移动目标防御，当前也不存在明确的定义，其目标是通过持续变换系统呈现在攻击者面前的攻击面，从而有效增加攻击者想要探测目标脆弱性的代价。2010 年发布的《网络安全游戏规则的研究与发展建议》中将移动目标防御的内涵描述为：期望能够创建、分析、评估和部署多样化、随时间持续变化的机制和策略，以增加攻击者实施攻击的复杂度和成本，降低系统脆弱性曝光和被攻击的概率，提高系统的弹性。实际上，移动目标技术是通过降低系统确定性、静态性和同构性来增加攻击复杂度从而防护一个系统的机制/方法的统称，通过变换系统配置（广义上的配置，包括硬件平台、软件版本、地址信息 、协议信息等）缩短系统某一配置属性信息的有效期，使得攻击者没有足够的时间对目标系统进行探测和对代码进行开发；或者同时降低其所收集信息的有效性，使其探测到的信息在攻击期间变得无效，以此

提高攻击者收集信息的代价和复杂性，降低系统被成功攻击的概率，提高系统的抗攻击能力。

移动目标防御与拟态安全防御分别是美国和中国所提出的用于改变网络安全对抗现状的革命性技术，均期望能从根本上改变当前网络“易攻难守”的局面。二者之间既有联系又有区别。

① 从整体目标来看，移动目标防御与拟态安全防御所期望达到的整体目标基本相同，均希望能够通过实施己方可控的变化来迷惑攻击者，从而增加攻击的难度和代价，有效降低攻击成功率。

② 从基本内涵/思想来看，拟态安全防御的基本思想是在功能等价条件下，以提供目标环境的动态性、非确定性、异构性、非持续性为目的，动态地构建网络、平台、环境、软件、数据等多样化的拟态环境，以防御者可控的方式在多样化环境间实施主动跳变或快速迁移，对攻击者则表现为难以观察和预测的目标环境变化，从而增大攻击难度和代价。从基本内涵/思想看来，拟态安全防御对其目标及手段描述得更为清晰，这在一定程度上对其适用范围及方法进行了限定；而移动目标防御的基本思想比拟态安全防御更为灵活和宽泛，若将该思想应用于被保护系统的具体属性上，则与拟态安全防御的思想相近。

③ 从应用场景来看，拟态安全防御是基于拟态计算的信息系统所具备的一种内在主动防御能力，也就是说，拟态安全防御依赖于拟态计算系统，是一种灵活的思想，可应用于某一系统之上，也可应用于具体方法中，其应用场景更广。

④ 从发展状态来看，拟态安全防御研究尚处于起步阶段，暂无其他公开发表的成果，而移动目标防御研究正处于快速发展期，已涌现出大量的研究成果。

11.5.2 移动目标防御技术的最新进展

近几年，移动目标防御技术已取得很多新进展，包括变形网络、自适应计算机网络、自清洗网络，以及移动目标 IPv6 防御、开放流随机主机转换技术等研究。

1. 变形网络

2012 年 8 月，美陆军授予雷声公司价值 310 万美元的“限制敌方侦察的变形网络设施”（MORPHINATOR）项目，为其研制具有“变形”能力的计算机网络原型机，该项目主要研究在敌方无法探测和预知的情况下，网络管理员有目的地对网络、主机和应用程序进行动态调整和配置，从而预防、延迟或制止网络攻击。

2. 自适应计算机网络

2012 年 5 月，美国堪萨斯州大学开始为美空军科学研究办公室研究“自适应计算机网络”，重点研究和量化移动目标防御对计算机网络的影响；将研究计算机网络通过自动改变自身设置和结构来对抗在线攻击的可行性，并开发有效的分析模型，以确定移动目标防御系统的有效性。

3. 自清洗网络

自清洗入侵容忍（SCIT）体系结构可以通过不断清洗服务器及不断变换个人服务器准

则来阻挡或限制网络攻击。这实际就是移动目标防御技术的一种应用，目前已取得多项研究成果。

4. 移动目标 IPv6 防御

移动目标 IPv6 防御（MT6D）提出了移动目标 IPv6 防御的新思路。它通过重复转动发送和接收者的地址来实现对用户隐私和目标网络的保护。研究结果表明，MT6D 不但是可行的，而且还能与新的 IPv6 地址无缝绑定。同时，MT6D 能够为平台和应用层提供一种有力的移动目标解决方案。

5. 开放流随机主机转换技术

开放流随机主机转换（OFRHM）技术就是利用开放流研究移动目标防御体系结构，实现 IP 地址的不可预知性及高速变换，同时保持配置的完整性，并最小化操作管理。研究结果表明 OFRHM 能够有效防御秘密扫描、蠕虫传播及其他基于扫描的攻击。

11.5.3　移动目标防御机制

当前已有大量具体的移动目标防御机制被提出，国外学者依据技术在执行栈中的位置层次将现有移动目标技术分为动态数据、动态软件、动态运行环境、动态平台及动态网络 5 个大类。下面主要介绍几种动态网络中基于通信参数变换的机制。

1. DYNAT

动态网络地址转换（Dynamic Network Address Translation，DYNAT ）通过变化报文头中的主机标识信息来防御网络嗅探攻击。在报文被路由之前，通过 DYNAT shim 对发送方添加到报文头中的初始地址信息（端口和地址）进行转换（转换算法依赖于一个预先设定好的随时间变化的参数），然后发送到公共网络中。DYNAT gateway 在接收到该报文之后，会对报头域进行逆转换来获得初始的身份信息，然后对其进行正常处理并发送给接收方。该技术能有效增加攻击者窃取有效信息以映射网络并发起攻击的难度，但是对用户不透明，且该机制被设计用来保护一组部署在集中式网关后面的静态节点，通过一个接口在被保护网络与外部网络之间为所有进入和出去的报文执行地址信息翻译工作，在网络配置动态性较高的情况下，可能会出现因无法通过一个集中的网关管理所有的通信和实现节点同步而失效的情况。

2. NASR

网络地址空间随机化（Network Address Space Randomization，NASR）通过在网络地址动态分配的环境中调节节点 IP 地址的更改频率来防御蠕虫攻击。该机制需要配置一个 DHCP 服务器在不同时间间隔内终止 DHCP 租约（lease）以实现地址随机化，地址的改变可以在节点重启时进行，也可以基于定时器设置来实现。该机制需要对 DHCP 服务器进行修改，且该机制必须部署在具有动态地址的网络中，因此部署代价比较高。此外，该机制实现的是局域网级别的地址随机化，因此所能提供的不可预测性较低。

3. IP 地址变化

Al – Shaer 等人提出了一系列的地址变换（IP address mutation）机制。首先提出的是一种应用于软件定义网（Software Defined Network，SDN）中的地址随机变换技术 OF – RHM（OpenFlow Random Host Mutation），通过 OpenFlow 控制器频繁地为主机分配随机虚拟 IP，且由 OF – switch 执行真实 IP 与虚拟 IP 之间的转换，使得网络中传输报文所带有的均为虚拟 IP，以增加攻击者对特定端主机进行探测的难度。鉴于 OF – RHM 在传统网络中难以部署，接着又提出了随机主机交换（Random Host Mutation，RHM）。RHM 的设计思想及实现方式与 OF – RHM 类似，主要区别在于虚拟 IP 分配方法和分配组件的不同：该机制使用低频变换（Low Frequency Mutation，LFM）和高频变换（High Frequency Mutation，HFM ）两种随机变换粒度来实现虚拟 IP 的分配，其中，一个低频变换间隔区间内包含多个高频变换间隔区间。在每一个低频变换间隔内，系统会为每一个主机选择一个满足多种限制条件的随机地址范围，而后在每一个高频变换间隔内，在上一个 LFM 间隔所分配的地址范围内随机选择一个虚拟 IP 分配给主机。虚拟地址的分配由移动目标控制器（Moving Target Controller，MTC）来执行，而真实 IP 与虚拟 IP 之间的转换则由移动目标网关（Moving Target Gateway，MTG）来实现。为了进一步提高变换机制所提供的安全度，又提出了时空地址变换（Spatio – Temporal Address Mutation）机制来动态变化“主机 – IP”绑定关系。在该机制中，与主机真实 IP 相对应的瞬时 IP（ephemeral IP，eIP）是以源标识和时间为参数来确定的，因此每一个瞬时 IP 仅能在特定时间间隔与另一个特定主机进行连接通信。与 OF – RHM 及 RHM 不同的是，在源网关和目的网关之间网络上传输的报文所带有的是源和目的主机的真实 IP，而不是虚拟 IP。这 3 个地址变换机制均能在保持主机真实 IP 地址不变的同时以较高的速率来改变主机的通信地址，因此能在保证对用户的透明性的同时使得攻击者所收集的信息快速失效，且普通用户并不知道通信对方的真实 IP 地址，从而无法通过常规手段收集有效信息，但是机制的实施复杂度较高。

4. MT6D

MT6D（Moving Target IPv6 Defense）是在 IPv6 下实现的网络层移动目标防御方法，用于对抗针对特定目标的窃听、攻击和主机追踪等。在该方法中，通信双方利用各自当前地址的接口标识符（Interface Identifier，IID）、一个共享的对称性密钥及系统时间，计算出下一步要使用的接口标识符并通告出去，然后使用新的接口标识符来进行通信。MT6D 可在一个会话期间进行多次网络地址变化而不会中断会话，因此对通信性能的影响不大。由于通信双方的地址持续在变化，且 IPv6 地址空间较大从而使得变化比较丰富，导致攻击者想要对通信双方发起攻击就必须耗费更多的资源和时间，从而增加了攻击的代价和困难度。现实中，MT6D 可应用于智能网格（smart grid）中保护对等端（peer）之间的通信，同时对对等端本身提供一定的保护以免遭攻击。但在同一时刻，路由器处要为同一个节点保存多个地址及相应的对应关系，因此增加了一定的存储开销。此外，MT6D 初始采用 UDP 作为其传输层协议，而 UDP 所创建的套接字仅能绑定在单个 IP 地址上，限制了 MT6D 在多地址情况下的适用性，因此可考虑将多宿主传输层协议，如 SCTP 应用到 MT6D 中，以改善其性能。

5. 端口跳变和（或）地址跳变（Port and/or Address Hopping）

端口跳变（port hopping）技术主要是对服务所使用的 UDP/TCP 端口号进行跳变，使得攻击者无法得知当前有效的服务端口号从而抵御 DoS/DDoS 攻击。Lee 等人提出采用一个以系统时间、服务器与用户之间的共享私钥为变量的跳变函数来进行 UDP/TCP 端口跳变，使得经过认证的用户因拥有密钥而能够得到服务器当前正在使用的端口号，而恶意用户无法得知当前的有效端口号。该技术与 UDP 和 TCP 协议相兼容，不需要改变现有协议，也不需要对 Internet 基础设施进行任何改变，可通过套接字通信来实现，易于实施，且能简化对恶意攻击报文的检测和过滤，但是该机制所采用的严格时间同步机制在延时和拥塞环境下适用性较低。网络地址跳变（network address hopping）的思路是通过跳变的方式使用多个被称为信道（channel）的数据连接（data connection）来传递一个通信会话的数据流，其中，信道是由目的 IP 端或〈源端，目的端〉元组来定义的逻辑通道，而一个通信端既可由 IP 地址来定义，也可由 IP 地址和端口号来共同定义，还可由应用标识符来定义。该机制改变了两个通信体之间的通信模式，能有效干扰和迷惑攻击者，且对用户应用完全透明，但是需要为通信的每一端都配置多个地址。

参考文献

[1] 李信满，赵大哲，赵宏，刘积仁．基于应用的高速网络入侵检测系统研究[J]．通信学报，2002，23（9）：1-7.

[2] 王英，向碧群．基于用户行为的入侵检测系统[J]．计算机工程，2008，24（9）：167-169.

[3] 蔡益朝，张维明，刘忠，等．基于遗传算法的实体分群问题的求解方法[J]．计算机工程，2007，33（5）：4-6.

[4] 马恒太，卿斯汉，等．基于 Agent 的分布式入侵检测系统模型[J]．软件学报，2000，11（10）：1312-1319.

[5] Paul S，Andrewsl，Jon T，Inspiration for the Next Generation of Artificial Immune System[C]．C. Jacob etal.（Eds.）：ICARIS 2005，LNCS 3627. 2005，126-138.

[6] Hofmeyr S A，Forrest S，Somayaji A. Intrusion Detection using Sequences of System Calls[J]．Joumal of Computer Security，2003，6（3）：151-180.

[7] Kephart J，Arnold W. Automatic Extraction of Computer Virus Signatures[C]．In proceedings of the 4th Virus Bulletin International Conference. Abingdon，1994，178-184.

[8] De Castro L N. Engineering applications of artificial immune systems[EB/OL]．Tutorial at ICARIS 2004，avaliable from http://artificial-immune-systems.org/ICARIS2004/icaris2004.htm.

[9] Bentley，Kim，Immune Memory in the Dynamic Clonal Selection Algorithm[C]．The 1st International Conference on Artificial Immune Systems（ECARIS-2002），University of kent at Canterbury，UK，2002（9）：113-118.

[10] Schultz M G，Eskin E，Zadok E，Bhattacharyya M，Stolfo S J. MEF：Malicious Email Filter，A UNLX mail filter that detects malicious windows executable[C]．In Proceedings of USENLX Annual Technical Conference，2001，245-252.

[11] Technica1 Report DCA-RT，Department of Computer Engineering and Industrial Automation，State University of Campinas，Brazil，[EB/OL]．2007：http://www.dcs.napier.ac.uk/~emmah/researeh_interests/AIS.htm.

[12] Arun Lakhoti，Eric Uday Kumar，and Michael Venable. A Method for Detectin Obfuscated Calls in Malicious Binaries[C]．IEEE Transactionson Software Engineering，2005，31（11）：955-968.

[13] 田畅，郑少仁．计算机病毒计算模型的研究[J]．计算机学报，2001，24（2）：158－163.
[14] Manadhata P. An attack surface metric[D]. Pittsburgh , Pennsylvania：Carnegie Mellon University，2008.
[15] Manadhata P. Game theoretic approaches to attack surface shifting[G]//Moving Target DefenseII：Application of Game Theory and Adversarial Modeling. Berlin：Springer，2013：1－13.
[16] Hobson T，Okhravi H，Bigelow D，et al. On the challenges of effective movement[C] //Proc of the 1st ACM Workshop on Moving Target Defense. New York：ACM，2014：41－50.
[17] Clark A，Sun K，Poovendran R. Effectiveness of IP address randomization in decoy－based moving target defense[C]//Proc of the 52nd Annual Conf on Decision and Control（CDC）. Piscataway，NJ：IEEE，2013：678－685.
[18] The WhiteHouse National Security Council. Cybersecurity progress after president Obama's address[EB/OL]. 2014. 03. 10，http://www. whitehouse. gov/administration/eop/nsc/cybersecurity/progressreports/july.
[19] NITRD Subcommittee. National cyber leap year summit 2009 co－chairs' report[EB/OL]. [2014. 06. 25]. https://www. nitrd. gov/nitrdgroups/index. php? title＝Category:National_Cyber_Leap_Year_Summit_2009.
[20] Okhravi H，Rabe M A，Mayberry TJ，et al. Survey of cyber moving target techniques，1166[R]. Lexington，Massachusetts：MIT Lincoln Laboratory，2013.
[21] Kewley D，Fink R，Low ry J，et al. Dynamic approaches to thw art adversary intelligence gathering[C] //Proc of IEEE DARPA Information Survivability Conf & amp；Exposition Ⅱ（DISCEX'01）. Piscataway，NJ：IEEE，2001：176－185.
[22] Antonatos S，Akritidis P，M arkatos E P，et al. Defending against hitlist worms using netw ork address space randomization[J]. Computer Netw orks，2007，51（12）：3471－3490.
[23] Lee H C J，T hing V L L. Port hopping for resilient networks[C]//Proc of the 60th Vehicular Technology Conf. Piscataway，NJ：IEEE，2004：3291－3295.
[24] Sifalakis M，Schmid S，Hutchison D. Network address hopping：A mechanism to enhance data protection for packet communications[C]//Proc of 2005 IEEE Int Conf on Communications. Piscataw ay，NJ：IEEE，2005：1518－1523.
[25] Jafarian J H H，Al－Shaer E，Duan Qi. Spatio－temporal address mutation for proactive cyber agility against sophisticated attackers [C] //Proc of the 1st ACM Workshop on Moving Target Defens. New York：ACM，2014：69－78.
[26] Jafarian J H，Al－Shaer E，Duan Qi. OpenFlow random host mutation：Transparent moving target defense using software defined networking[C]//Proc of the 1st Workshop on Hot Topics in Software Defined Networks. New York：ACM，2012：127－132.
[27] Al－Shaer E，Duan Qi，Jafarian J H. Random host mutation for moving target defense[G]//Security and Privacy in Communication Networks. Berlin：Springer，2013：310－327.
[28] 蔡桂林，王宝生，王天佐，等．移动目标防御技术研究进展[J]，计算机研究与发展，2016，53（5）：968－987.
[29] BassT. Multisensor data fusion for next generation distributed intrusion detection systems. In：Proc. of the'99IRIS National Symp. on Sensor and Data Fusion. Laurel，1999：24－27.
[30] Bass T. Intrusion systems and multisensor data fusion. Communications of the ACM，2000，43（4）：99－105..
[31] Wang J，Zhang FL，Fu C，Chen LS. Study on index system in network situation awareness. Computer Applications，2007，27（8）：1907－1909.
[32] Ticha B，Ranchin T. A case based reasoning data fusion scheme：Application to offshore wind energy resource mapping. In：Proc. of the Int'l Conf. on Information Fusion（FUSION），2006：1－5.
[33] Gad A，Farooq M. Data fusion architecture for maritime surveillance. In：Proc. of the Int'l Society on Information Fusion（ISIF），2002：448－455.

[34] Kadar I. Knowledge representation issues in perceptual reasoning managed situation assessment. In：Proc. of the FUSION，2005：13－15.
[35] Hall D，Llinas J. An introduction to multisensor data fusion. Proceedings of the IEEE，1997，85（1）：6－23.
[36] Blasch E，Piano S. JDL level 5 fusion model "user refinement" issues and applications in group tracking. In：Proc. of the Signal Processing，Sensor Fusion，and Target Recognition XI，SPIE Vol. 4729，2002：270－279.
[37] Blasch E，Piano S. DFIG level 5 issues supporting situational assessment reasoning. In：Proc. of the FUSION，2005：35－43.
[38] Endsley M. Situation awareness global assessment technique（SAGAT）. In：Proc. of the IEEE'88 National Aerospace and Electronics Conf.（NAECON），1988：789－795.
[39] Tadda G，Salerno J，Boulwarea D，Hinmana M，Gorton S. Realizing situation awareness in a cyber environment. In：Multisensor BV，ed. Proc. of the Multisource Information Fusion，SPIE Vol. 6242，2006：1－8.
[40] Zhuo Y，Zhang Q，Gong ZH. Cyberspace situation representation based on niche theory. In：Proc. of the ICIA. Zhangjiajie，2008：1400－1405.
[41] Klir G，Yuan B. Fuzzy Sets and Fuzzy Logic. New York：Prentice Hall，1995.
[42] Chert LY，Huang J. Survey of research on measure of uncertainty. Journal of Circuits and Systems.
[43] Genon P，Smith B. SNAP and SPAN：Towards dynamic spatial ontology. In：Proc. of the Spatial Cognition and Computation，2003：137－171.
[44] Little E，Rogova G. Ontology meta－model for building a situational picture of catastrophic events. In：Proc. of the FUSION，2005：796－803.
[45] Chen XZ，Zheng QH，Guan XH，Lin CG. Quantitative hierarchical threat evaluation model for network security. Journal of Software，2006，17（4）：885－897.
[46] Cai W. Extension Engineering Method. Beijing：Science Press，1997.
[47] 龚正虎，卓莹，网络态势感知研究[J]，软件学报，2010，21（7）：1605－1619.

第 12 章　物联网应用层云安全技术

云计算和物联网的结合其实是物联网迅速发展的需要。智慧物联网模型能够自动进行数据处理和交换操作，在能保证计算能力和存储能力十分高效的同时，还可以保证数据安全。除此之外，还需要具备明显的成本优势。云计算技术恰恰满足了上述许多的优势，因此物联网技术的彻底有效的实施和应用必然不能缺少云计算。本章在分析物联网与云计算融合趋势的同时，重点介绍云计算面临的安全问题及其安全关键技术。

12.1　云计算简介

12.1.1　云计算的概念

由于云计算是由不同的企业和研究机构同步推进的技术，所以关于云计算的定义有很多，至今并没有一个公认的定义和标准。比较典型的定义有如下几种。

国外学者 Ian Foster 定义云计算为一个由规模经济驱动的大型分布式计算模型，在该模型中，抽象、虚拟化、动态可伸缩并可管理的计算资源、存储资源、平台和服务构成了一个资源池。资源池中的资源通过互联网，按需提供给池外的用户。

文献“A break in the clouds：towards a cloud definition”归纳的云计算定义为：云是由易于使用的虚拟资源构成的一个巨大资源池，包括硬件资源、部署平台及相应的服务。根据不同的负载，这些资源可以动态地重新配置，以达到一个最理想的资源使用状态。资源池中的资源是按需付费的，服务提供商通过服务等级协议（Service Level Agreement，SLA）保证用户的服务质量。

维基百科上将云计算定义为：云计算是一种以互联网为载体的计算模式，在这种模式中，共享的软、硬件等资源可以按需地提供给计算机或其他设备。云计算拥有如同电网一般的运行方式，将计算、存储等服务如同水、电、煤气一般提供给需要的消费者。

综合以上定义，可以将云计算归纳为：云计算以虚拟化技术为核心，虚拟化技术将共享的硬件和软件资源抽象化成一个统一的资源池，通过互联网这个载体，向用户按需地提供所需的资源。其特点在于多用户共享、大数据处理与大数据存储。

12.1.2　云计算的特点

从云计算的概念中可以得出云计算具有大规模、多用户、虚拟化、高可靠性、可伸缩性、按需服务、成本低廉七大特性。其中可伸缩性、按需服务是区别于传统 IT 服务的新特性。

1. 大规模

云计算诞生之初的使命就是使用大量的机器，解决大数据集处理的问题，并存储海量数

据。因此，不管是从云计算本身的规模，还是其所处理数据的规模，云计算的一大主要特性就是大规模。Google 的云计算平台早已经拥有上百万台服务器，其他 IT 巨头的服务器数量也在百万级别的规模上。

2. 多用户

云计算的另一大特性就是将资源分享给许多用户同时使用，不同于以往的计算模式将特定的资源交付给特定的用户使用。多用户不仅体现在用户数量上，而且体现在用户种类上，不同服务需求的用户也可以同时使用同一个云平台，云平台使用虚拟化技术满足不同的用户需求。

3. 虚拟化

虚拟化技术是云计算的核心，云计算的虚拟化体现在两个方面：第一种虚拟化是将多台服务器虚拟化成统一的资源提供给用户，使用户可以透明地使用所需的资源而不需要关心底层的技术实现细节；另一种虚拟化则是将一个物理资源虚拟成多个虚拟主机提供给不同的用户使用，通过虚拟化来隔绝用户。通过虚拟化技术，云计算可以将资源统一成资源池，动态地提供给用户。

4. 高可靠性

云计算具有高可靠性，体现在云计算的容错机制上。云计算的容错机制在存储上表现为多副本容错，即将同一个内容备份多个副本，通过跨机架存储，甚至是两地三中心模式来保证数据的可用性。而在计算服务上，高可靠性主要表现为计算节点的可互换，当某个节点发生故障不能提供服务时，由另一个节点及时取代故障节点继续提供服务，从而实现服务的不中断。

5. 可伸缩性

云计算的规模是可以动态伸缩的，云计算服务提供商可以根据用户的规模来决定提供给用户的资源，并根据用户的增长速度动态地扩充自己的云系统。

6. 按需服务

云计算的按需服务特性主要针对的是云计算用户，区别于传统的 IT 服务模式，用户可以根据自己的真实需求使用云资源，而不需要考虑未来需求的增长。

7. 成本低廉

云计算是一种低成本的服务，对于云计算服务提供商来说，他们可以使用廉价的服务器构建高可靠性的云平台。而对于用户来说，则可以按需使用廉价的服务。

12.1.3　云计算的分类

云计算按服务模式可以分为通信即服务（Communication as a Service，CaaS）、硬件即服务（Hardware as a Service，HaaS）、基础设施即服务（Infrastructure as a Service，IaaS）、平台即服务（Platform as a Service，PaaS）和软件即服务（Software as a Service，SaaS）。现在

主流的分类是将云计算从下层到上层分为 IaaS、PaaS 和 SaaS。根据部署模式，云计算还可以分成公有云、私有云及混合云。从服务模式和部署模式两方面可以构建出一个云计算的分类模型，该模型如图 12-1 所示。

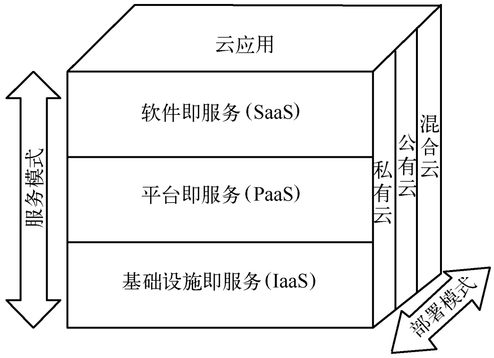


图 12-1　云计算分类模型

IaaS，顾名思义，就是将硬件等基础资源作为服务直接提供给用户。用户可以在这些基础设施上直接部署自己的操作系统镜像，不管是使用 Windows 系统还是 Linux 系统。IaaS 最大的特点就是用户可以直接接触云平台的底层资源，因此其使用的自由度很大。IaaS 主要包括四种类别的服务：资源管理服务、虚拟机服务、克隆服务和安全服务。

PaaS 是一种相对较新的服务模式，在云计算早期并没有 PaaS 的概念，几乎所有的云应用都是直接在云平台的硬件层上开发、运行的。随着云服务模型的明晰，为了方便开发和部署 SaaS，提出了 PaaS 的概念。PaaS 最大的特点是用户不需要管理和控制云基础设施，用户只需要使用平台提供的 API 就可以很方便地开发和部署云应用。

SaaS 是一种软件交付方式，也是用户最为熟知的云计算服务方式，其主要的形式是用户通过瘦客户端，最常使用的是浏览器，直接使用云服务提供商提供的服务，并将自己的数据存储在云端。

公有云，即广义的云计算，由第三方云服务提供商提供用户所需的所有资源。用户不拥有云系统的任何资源，同时用户也失去了对自己数据的所有权。公有云计算虽然成本低廉，使用方便，但同时也存在安全问题，主要包括对服务提供商的信任问题、来自于恶意内部用户的攻击问题等。

私有云，主要指由企业的 IT 部门自行配置的云系统，用户拥有私有云系统的所有权。虽然构建私有云的成本比较高，但是由于系统相对封闭，用户可以绝对控制云系统，所以私有云具有很高的安全性。

混合云，即将私有云与公有云混合使用，也包含混合使用不同云计算提供商的云系统。混合云同时包含了私有云和公有云的优点，用户可以将重要的数据保存在私有云中，当需要公有云服务时，再提交给公有云，使用结束即收回。混合云是现在云计算研究的热点。

12.2　物联网与云计算的融合

12.2.1　与云计算相融合是发展必然

物联网的发展离不开云计算技术，主要有以下几部分原因。

1. 成本优势明显

物联网技术中，IT 基础设施的成本问题一直是最具有竞争力的。信息技术产业的开销在业界内被主要划分成三部分，即硬件开销、能耗费用及管理成本。根据市场调查，近十五年来左右 IT 设施的开销之中硬件开销基本持平，提升迅速的开销主要是能耗费用和管理成本两类。云计算在资源的利用率方面比传统的互联网数据中心有很大优势，传统的互联网采用的技术手段相对来说比较低端，在互联网中访问量会因为时间的不同而不同，这就使得很大一部分资源被空缺，从而导致资源的平均利用率很低。恰恰相反的是，高效的云计算平台提供给用户的是一种可伸缩的弹性服务端，它可以根据不同用户的需求来分配和释放资源，另外，云计算平台内规模庞大，数据用户众多，因此其平均利用率提高很多，从这个层面上分析，云计算技术给物联网的实现大大降低了运营成本。

2. 核心技术优势明显

物联网所构建的智能化网络体系给人类生产生活都带来了巨大的便利。由于其功能的庞大，所要求的技术手段也需要非常先进。物联网包含众多的使用对象和数据、数以亿计的节点，这其中就包括了众多的大型网络服务器和超级集群，因此物联网的实施必须依靠发达的计算和通信技术。物联网之中庞大的数据产生和收集的过程具有实时性和不间断性等特点，时间的延长也必然会导致数据量的扩大。这些现象都将要求物联网技术中数据挖掘技术的提高。由于数据量大、节点数目有限、数据存储地点不同、数据安全性要求高等原因，导致物联网技术的数据挖掘不仅仅是传统意义上的数据统计分析，而是一种潜在模式的挖掘技术。云计算技术最主要的应用是分布式技术，它很好地解决了上述问题，可以有效地管理和控制多模式、多源、多位置的不同数据，并且保证了数据的安全性，所以说云计算的技术手段也是物联网不可或缺的。

3. 计算能力和存储能力有优势

云计算相对于传统的计算模式具有高速互联网连接，近乎无限的存储和计算能力等优势。云计算主要通过网格计算等技术把多个计算机实体结合成为一个功能强大的计算系统，并且通过相关技术把超强的计算能力均匀分配到终端用户之中。云计算的最大优点是能够像超级计算机一样对应用进行并行处理，而且耗费很低的成本。除了计算能力外，存储能力的优势也在云技术之中表现明显。人们在近几年所产生和使用的数据都是非结构化的数据，越来越多的用户愿意把自己的信息和资源传到云中。因此，云存储应该具有高可靠性、高可用性和自动容错能力等特点。

12.2.2　基于云计算的物联网系统

物联网在应用过程中呈现出诸多云计算特征，如对资源的大规模和海量需求、资源负载变化大、以服务方式提供计算能力等，从而适合采用云计算技术建立物联网应用系统。基于云计算的物联网应用系统主要由云基础设施、云平台、云应用和云管理四部分组成。

1. 云基础设施

云基础设施是指通过物理资源虚拟化技术，使得平台上运行的不同行业应用及同一行业

应用的不同客户间的资源（存储器、CPU 等）实现共享，并提供资源需求的弹性伸缩；通过服务器集群技术，将一组服务器关联起来，使其在外界看起来如同一台服务器，从而改善平台的整体性能。

2. 云平台

云平台是物联网运营云平台的核心，实现网络节点的配置和控制、信息的采集和计算。可以采用分布式存储、分布式计算基数实现对海量数据的分析处理，以满足大数据量且实时性要求非常高的数据处理要求。

3. 云应用

云应用用于实现行业应用的业务流程，可以作为物联网运营云平台的一部分，也可以集成第三方行业应用，但在技术上应通过应用虚拟化技术，实现多租户，让一个物联网行业应用的多个不同租户共享存储、计算能力等资源，提高资源利用率，降低运营成本，在共享资源的同时又相互隔离，保证用户数据的安全性。

4. 云管理

云管理采用了弹性资源伸缩机制，用户占用的电信运营商资源随时间在不断变化，需要平台提供按需计费的支持能力。

12.2.3 云计算与物联网的融合模式

云计算与物联网各自具备很多优势，将两者有机融合将具有更高的应用效益。云计算与物联网的融合可以采用以下三种模式。

1. 单中心多终端模式

此类模式分布在范围较小的各物联网终端（传感器、摄像头或手机等），把云中心或部分云中心作为数据/处理中心，终端获得信息、数据由云中心统一进行处理及存储，云中心提供统一界面给使用者操作或查看。这类应用非常多，如小区及家庭监控、对某一高速路段的检测、对幼儿园小朋友的监管及某些公共设施的保护等。

2. 多中心大量终端模式

对于很多区域跨度加大的企业、单位而言，该模式比较适合。例如，一个跨多地区或多国家的企业，因其公司较多，要对各公司或工厂的生产流程进行监控、对相关的产品进行质量跟踪等。

3. 信息、应用分层处理的海量终端模式

这种模式可以针对用户范围广、信息及数据种类多、安全性要求高等特征打造。当前，客户对各种海量数据处理需求越来越多，应根据客户需求及云中心的分布进行合理的资源分配。

12.3　云计算安全问题

云计算作为一种新的服务和计算模式，本身并没有脱离传统信息安全概念覆盖的范围，依然面临着各种传统安全威胁，包括物理实体的安全、网络的安全、主机的安全等。传统的安全技术，尤其是一些基础性技术仍然适用于云计算系统，如加解密技术、入侵检测技术和防火墙技术。但由于本身的特点又会导致云计算产生一些新的安全问题，如内部虚拟机相互攻击、超级管理员权限、多租户安全隔离等。

云计算的服务模式从底层的 IaaS 模式到顶层的 SaaS 模式会向用户提供不同的服务和接口，用户也会接触到不同的资源，这就决定了在不同服务层会有不同的安全问题。

12.3.1　IaaS 安全问题

IaaS 会提供给用户所有设施的使用权，使用户可以自由地部署自己的操作系统镜像，因此 IaaS 模式面临的安全问题是最多的，而且这些问题也主要与云平台本身相关。IaaS 层包括的主要安全问题如下。

1. 物理设施安全问题

物理设施安全包括服务器、通信网络、电源等物理实体的安全。基础物理实体的安全是整个云计算系统安全的大前提。虽然云计算并没有改变传统物理实体所面临的安全问题，如自然灾害、管理人员的误操作等，但是因为云平台，尤其是公有云平台托管的用户将比传统 IT 系统多很多，所以物理设备的安全问题会造成更广泛的影响。

2. 数据传输安全问题

数据传输安全主要是指如何确保数据传输过程中的保密性和完整性。在云环境中该问题又分为用户与云平台数据传输时的安全问题和云平台内部数据传输的安全问题。用户与云或是不同云系统之间的数据传输一般是加密的，并采用 FTPS、HTTPS 等安全传输协议。但是云系统内部的数据传输一般是非加密的，因为现有的系统还无法直接处理加密数据，这是云中数据传输主要面临的安全问题。

3. 数据存储安全问题

IaaS 层的数据存储安全涉及数据的保密性、完整性和可用性三个方面。数据的保密性主要是通过身份验证、访问管理和静态数据加密保证，完整性主要通过访问管理和数据校验等方式进行保证，而目前数据的可用性主要通过冗余备份的方式来保证。数据存储的安全问题还包含一个经常被忽视的问题——数据残留问题。如何保证用户删除的数据在云中被彻底销毁，不会被服务提供商或其他用户恶意恢复，这是目前一直没有受到重视的问题。目前绝大部分云服务提供商都没有关注这一严重的问题，随着越来越多的用户将自己的数据，甚至是敏感数据迁移到云平台上，如何解决这一问题也变得越来越紧迫。

4. 虚拟化安全问题

虚拟化安全问题是 IaaS 层主要的安全问题，现在用户使用的 IaaS 服务基本上都是基于

虚拟主机的服务。虚拟化安全问题主要包括两个方面：虚拟化软件本身的安全问题和虚拟主机的安全问题。虚拟化软件位于裸机与用户实例之间，提供用户创建与删除虚拟化实例的能力。不安全的虚拟化软件，会直接造成用户实例的不安全，如可能造成用户实例的非授权访问、用户实例的非法删除等。虚拟主机的安全问题与传统主机的安全问题基本相同，主要面临的威胁有用户劫持、攻击没有防火墙的虚拟主机、攻击未安补丁的主机服务漏洞等。同时，虚拟主机也有自身特有的安全问题，如内嵌在虚拟镜像组件中的木马等。

5. 接口与 API 安全问题

IaaS 模式主要是通过接口或 API 向用户提供特定的服务，如资源管理、系统监控等服务。但是当前 IaaS 的接口和 API 本身还不够安全，尤其像是提供认证服务、加密服务、访问控制服务的接口必须能够保证不会被偶然或恶意地规避，从而导致整个系统安全防护的丧失。不安全的 API 也会造成恶意用户利用接口对系统内或系统外进行非法攻击，以及对云资源的滥用。造成不安全 API 的原因有很多，如匿名访问、重用的令牌或密码、认证与数据传以明文的方式进行、不灵活的访问控制和认证，以及有限的监控和日志能力等。

6. 共享技术带来的安全问题

IaaS 层准许用户直接使用底层的基础架构，却没有提供不同用户之间的有效隔绝措施。虽然一些云服务提供商在用户的操作系统与底层架构之间设置了一层管理程序，以保证用户不会不恰当地使用底层服务，但是这样的措施还不能完全保证用户之间的隔绝。如果不能有效地管理用户、隔绝用户，可能会造成用户对共享资源的非法使用，甚至是内部用户之间的互相攻击。

7. 应用程序安全问题

IaaS 层上的用户应用程序主要是指用户自己部署的操作系统镜像。在实际应用中，提供 IaaS 服务的云服务提供商会将用户的操作系统实例当作黑盒来处理，这主要源于用户在 IaaS 层上具有比较大的自由度，因此云服务提供商对于用户应用程序的运行情况并不了解，相应的应用程序安全措施必须由用户自己提供。而在 PaaS 和 SaaS 层，用户的应用程序则会得到云服务提供商一定的安全保证。因为云服务提供商无法了解用户应用程序的安全性，如何及时发现、处理和隔离不安全的应用程序是云服务提供商需要解决的主要安全问题。

12.3.2 PaaS 安全问题

PaaS 层所面临的安全问题主要有：API 的安全问题、模块整合安全问题、数据的安全问题。

1. API 的安全问题

PaaS 层的特点就是提供给用户丰富的 API，让用户可以通过这些 API 在云平台上开发、部署自己的应用。与 IaaS 层面临同样的安全问题，不安全的 API 会直接导致用户开发的应用程序安全性降低。同时，PaaS 层也有其特有的 API 安全问题——各个云服务提供商没有

统一的 API，也就是说，每个云平台能提供的安全保障是不一样的，这会导致在某一个云平台上安全的应用程序移植到另一个云平台后变得不再安全。例如，GAE 使用 Python 或 Java 对象设定用户安全配置，而 Force. com 使用专有的 Apex 语言 API 来设定安全参数，不仅是使用的编程语言不同，两者能提供的安全服务水平也不同。

当下 PaaS 提供的安全保障 API 数量还不足，只能向用户提供诸如 SSL 配置、基本的访问控制和权限管理等基本安全功能。因此，增加安全功能 API 的数量，并向用户提供完善的安全保障服务也是现在 PaaS 服务所要考虑的一个主要问题。

2. 模块整合安全问题

PaaS 层将一些底层的服务集合为一个整体，然后以一个统一的 API 向用户提供相应的服务，这与用户在 IaaS 层使用 API 和接口的方式极为不同。在 IaaS 层，用户直接通过接口与底层的服务通信，这些服务相对都是独立、单一的服务，不存在服务整合的问题。由于缺少相关的标准，这种集成后的服务模块的安全性无法保证，同时也难以评估这些模块的安全性，用户不得不针对某一 API 进行大量的测试，以保证其提供的模块是安全可靠的。

3. 数据的安全问题

PaaS 层的数据安全问题主要来源于应用程序使用的静态数据是不加密的，因为加密后的数据将无法被索引和查询。但是这些不加密的静态数据很有可能被来自内部的攻击者或非授权访问者窃取，从而破坏数据的保密性。

12. 3. 3 SaaS 安全问题

SaaS 的特性决定了云服务提供商对整个服务的安全性负主要责任，因为用户只是使用服务提供商提供的应用程序，除了避免误操作外，用户只能使用服务提供商提供的安全措施。SaaS 层主要面临的安全问题有权限管理的安全问题、数据处理的安全问题和云服务不透明问题。

1. 权限管理的安全问题

在 SaaS 层中，服务提供商提供的身份验证与访问控制等权限管理手段可以说是用户可用的唯一安全控制手段。但是现在 SaaS 层提供的权限管理手段还没有实现先进的细粒度访问控制，并且存在访问控制标准不一致问题，如 Google Docs 处理文件的内嵌图片的访问控制机制与旧版本文件的访问控制存在不一致问题，这就可能造成当用户停止共享这个文件时，其他用户仍然可以共享文件中的内嵌图片。

2. 数据处理的安全问题

与 PaaS 层遇到的问题相同，SaaS 层的应用程序在使用静态数据时也不会对数据进行加密，存在数据泄露的风险。但是 SaaS 层数据泄露的风险比 PaaS 层要大，这是因为在 PaaS 层用户可以通过自己开发一定的防范程序来预防数据泄露，而在 SaaS 层用户无能为力，所有的安全防范只能由云服务提供商提供。

3. 云服务不透明问题

IaaS 和 PaaS 服务模式会向用户开放相应的服务接口，用户可以借此评估整个云平台的安全性，尤其是在 IaaS 层，用户的自由度更大。但在 SaaS 层，用户直接使用服务，无法评估整个云平台的安全性，这会使用户的数据受到未知安全风险的威胁。用户和云服务提供商应该在保密协议的约束下，分享包括结构、实现、部署、白盒和黑盒测试等方面的信息，以使用户可以有效评估云平台安全性。当下，云服务提供商，尤其是 SaaS 服务提供商不愿意过多提供云平台的技术细节等安全评估内容。

12.3.4 其他安全问题

除了上节列出的云服务模型中每层所面临的具有针对性的安全问题外，云计算还面临着以下安全问题。

1. 恶意内部使用者问题

来自于内部恶意用户的威胁是信息系统的传统威胁，但是由于云计算将多用户聚集在一个管理域中，用户共享一个平台，使得这种威胁变得更加严重。同时，云计算服务提供商缺乏必要的招聘和培训手段，并且没有相应的授权与监控手段，云计算服务提供商无法保证其员工不做出侵犯用户利益的行为，如窃取数据、破坏服务等。

2. 账户/服务和传输的劫持问题

账户或服务劫持也不是一种针对云计算的新威胁，但是这种威胁在云平台上会产生更严重的后果。当攻击者劫持了云用户的账户，攻击者不仅能获得云用户的数据，监视用户在云平台上的活动，同时也可以借助被劫持者的账户攻击云平台中的其他用户，这样就使危害得以成倍放大。

3. 云计算的滥用问题

IaaS 和 PaaS 模式为用户提供了几乎无限的计算、网络和存储资源，可以说只要用户拥有足够的金钱为使用这些资源付费，用户就可以立即使用这些资源。由于云服务提供商缺乏必要的审查与监管，一些恶意用户可以使用这些资源进行一些违法活动，如暴力破解密码、将云平台作为发动分布式拒绝服务（Distributed Denial of Service，DDoS）攻击的源头、利用云计算控制僵尸网络和托管非法数据等。

4. 来源于未知风险的安全问题

云计算部署模式中的公有云和混合云主要的服务模式就是托管用户的数据和服务，虽然这种模式使用户不用关心技术细节，不需要维护基础设施，但是也使用户失去了对数据和服务的所有权。不清晰的第三方平台使用户无法有效评估自己的数据和服务的安全性，也使云系统充满了未知的风险。这种安全问题在 SaaS 服务模式中更为突出。

5. 服务可用性问题

服务可用性问题是云计算的一个核心安全问题，云中托管的数据和服务来源于数量庞大

的用户群，如果云平台发生服务不可用问题，造成的影响将会成倍放大，远超出传统 IT 系统。造成服务中断的威胁既来源于云系统内部，也来源于其外部。内部的威胁主要是云平台自身的可靠性问题，如发生服务器宕机、数据大规模丢失等都会造成云服务不可用。云平台的可靠性可以通过容灾备份技术得以加强。服务可用性的外部威胁主要是 DDoS 攻击威胁。

DDoS 是现今网络上主要的攻击方式之一，攻击者利用所控制的僵尸网络等网络资源向目标主机发送攻击包，以达到阻止目标服务器正常提供网络服务的目的。云计算平台会托管众多不同的网络服务，必然会吸引攻击者向云平台发动 DDoS 攻击。由于云计算的多租户特性，DDoS 所造成的影响将超过对传统服务器攻击造成的影响。因此，如何防范针对云平台的 DDoS 攻击，以及如何利用云平台的资源优势识别、阻止 DDoS 攻击是现在的一个研究热点。

根据云计算的分类模型可知，云计算除了有三层服务模式外，也有三种主要的部署模式：私有云、公有云和混合云。不同的部署模式也会面临不同的安全问题。对于私有云来说，用户不用担心这种新模式带来的新的安全问题，虽然用户的 IT 架构可能会随着私有云的部署发生变化，但是这个系统的网络拓扑不会有明显的变化，用户可以完全掌握私有云的安全边界。而在公有云和混合云中，由于用户会将服务和数据托管到第三方云平台，就会面临前面两节总结的安全问题。将云计算的分类模型与云计算安全问题结合在一起，构建出云安全问题立方模型，如图 12-2 所示。

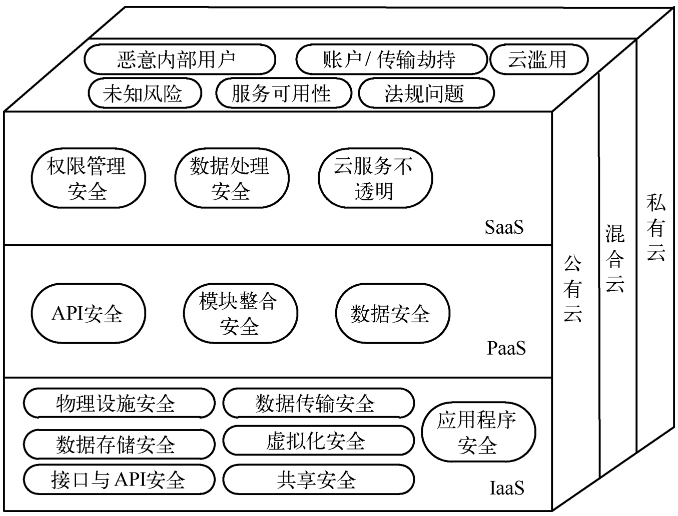


图 12-2　云安全问题立方模型

12.4　云安全关键技术

1. 容灾与恢复技术

容灾与恢复技术主要解决的是云计算服务可用性的问题，容灾与恢复技术可以帮助云系统在自然灾害、系统故障、人为失误中快速恢复丢失的数据和间断的服务，以提高云计算服务的可靠性。现在主要的容灾与备份解决方案如下。

1）HDFS 的冗余存储方案

Apache 的 Hadoop 系统是一个在计算机集群上使用简单编程模型来处理大数据集的软件框架。Hadoop 系统将计算与存储功能从单一服务器扩展到了一个包含数千台服务器的集群上，其一大特点就是将本地存储与本地计算相结合，集群中的每个节点只处理存储在自己本地硬盘上的数据。

Hadoop Distributed File System（HDFS）是一种运行于商业硬件上的分布式文件系统。它与现在流行的一些分布式文件系统有很多相似的地方。而其与其他系统的区别之处又显得极其重要。HDFS 提供了对应用数据的高吞吐量能力，适合大数据集应用场景。

HDFS 的 NameNode 的元数据和 DataNode 的 Block 都有冗余存储。NameNode 一般采用双机备份策略，DataNode 则是在不同的 DataNode 中存储同一个 Block 的三个副本。NameNode 的元数据在以前的版本中使用 Secondary NameNode 进行数据备份，但在新版本中则采取 Checkpoint Node 或 Backup Node 方案备份元数据。NameNode 使用 fsimage 和 edits 这两个文件持久化自己的命名空间，其中 fsimage 存储的是最近一次命名空间的备份，edits 则是针对最近备份点的所有改动的日志文件。当一个 NameNode 启动时，将合并 fsimage 和 edits 这两个文件以得到最新的文件系统元数据，然后覆盖原来的 fsimage 并开始新的 eidts。Checkpoint Node 的作用是周期性地从运行中的 NameNode 上下载 fsimage 和 edits，再合并这两个文件，最后将最新的文件系统镜像上传到 NameNode 上。基于此，当一个 NameNode 节点因为某个原因宕机后，Checkpoint Node 能及时提供最新的文件系统镜像给新启动的 NameNode，从而保证用户数据不丢失和服务的连续性。Backup Node 则是更为高效的备份方式，Backup Node 在内存中备份最新的 NameNode 文件系统镜像，并持续地接受来自 NameNode 的 edits 文件以更新文件系统镜像，由于是在内存中执行备份，所以 Bakup Node 的效率和可靠性比 Checkpoint Node 要高。现在的 HDFS 版本只支持一个 Backup Node，在以后的版本中会加入对多 Backup Node 的支持。

在 Datanode 上则是将一个 Block 的三个不同副本通过机架感知的方式存放到不同的 DataNode 上，其中两个副本存储在同一个机架上的两个不同 DataNode 中，另一个副本在另一个机架上，这样就可以预防机架整体失效的问题，并且可以保证恢复时的效率。HDFS 还采用负载均衡机制，让 DataNode 均匀存储 Block，这样可以防止存储集中产生的不可靠问题。

2）异地灾难备份机制

HDFS 的本地冗余备份可以解决一般的故障恢复问题。机架感知技术虽然解决了跨机架的安全问题，但是面对如大地震这样的大范围灾难时却无能为力。像银行、政府这样需要高安全性的机构，需要解决灾后服务恢复的问题，因此像银联数据中心这样的金融机构采用了“两地两中心”或“两地三中心”这样的异地灾难备份机制。两地三中心架构如图 12-3 所示。

2. 身份认证与访问管理技术

身份认证和访问管理（Identity and Access Management，IAM）是一套业务处理流程，是一个支持数字身份管理与资源访问管理、审计的基础结构。IAM 的关键功能有 SSO、认证管理、基于策略的集中式授权和审计、动态授权等。一个合理的 IAM 架构可以有效地帮助系统实现资源的安全管理和使用，以保证资源的保密性及不被非法使用。IAM 目前对云计算最有益的实践主要在三个方面：认证、授权和审计，见表 12-1。

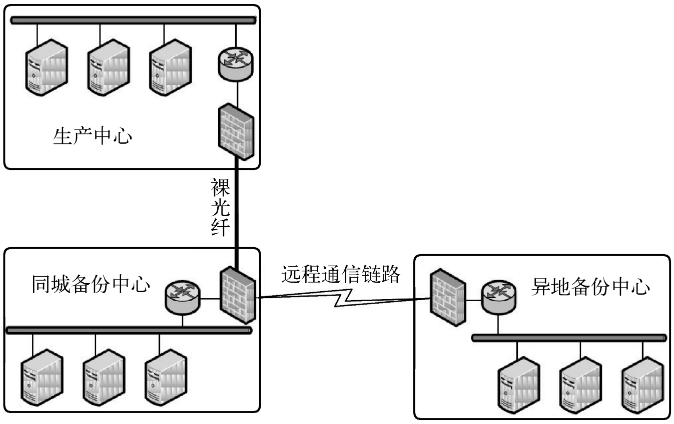


图 12-3　两地三中心架构图

表 12-1　云计算中认证、授权和审计的功能

认证	云计算的认证包括确认接入云系统的用户或系统的身份
授权	一旦认证成功，授权给用户使用相应云资源的权限
审计	审计的过程是检查认证和授权的记录，判断是否遵循了设定的规则，审计也可以帮助检测云系统的漏洞

IAM 在云中主要是通过一些标准和协议进行身份管理，包括安全断言标记语言（Security Assertion Markup Language，SAML）和开放认证（Open the Authentication，OAuth）协议。SAML 是基于 XML 语言的一种可以加强两个实体间认证和授权属性的语言，在云计算中则应用于云服务提供商和身份验证服务提供商之间。SAML 语言的主要目标是通过网络实现用户的 SSO，同时 SAML 还支持数字签名和加密。图 12-4 展示了 SAML 在云服务提供商、用户和身份验证服务提供商之间的工作流程。

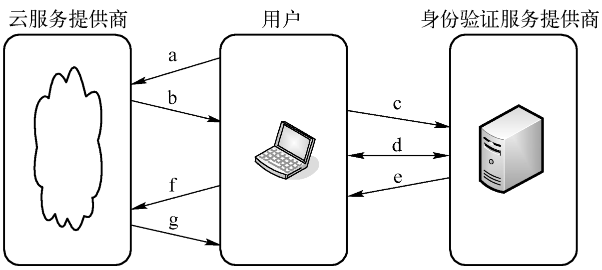


图 12-4　SAML 的工作流程

在图 12-4 中：

（a）用户从云服务提供商请求一个云应用；

（b）云服务提供商将用户重定向到身份验证服务提供商的 SSO 网站；

（c）浏览器登录 SSO 网站；

（d）在用户和身份验证服务提供商之间交换认证；

（e）身份验证服务提供商通过编码后的 SAML 语言回复用户；

（f）用户将 SAML 响应发送给云服务提供商；

（g）云服务提供商验证后，准许用户使用自己的云应用。

相对于 SAML 解决了云计算中用户登录的问题，OAuth 则准许用户在不暴露自己个人信息的情况下访问别的服务提供商提供的服务和数据。图 12-5 展示了 Google 的 OAuth 使用实例。

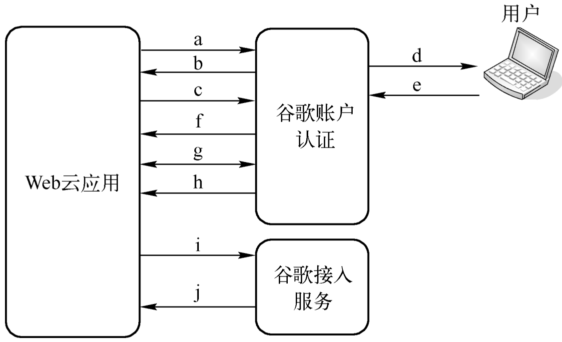


图 12-5　Google 的 OAuth 使用实例

在图 12-5 中：

（a）云应用向 Google 授权系统请求对 OAuth 请求令牌的授权；

（b）Google 会向云应用返回没有授权的请求令牌；

（c）云应用引领用户转到 Google 的授权页面去申请授权令牌；

（d）用户在 Google 的授权页面认证后，会被询问是否准许该云应用访问用户自己的数据；

（e）如果用户不准许云应用访问自己在 Google 上的数据，则会被转到 Google 的一个页面而不是云应用的页面；

（f）如果用户准许该云应用程序访问自己在 Google 上的数据，用户将被转向该应用的页面，该应用取得授权后的请求令牌；

（g）授权后的请求令牌将会在云应用与 Google 认证授权系统之间进行交换；

（h）当 Google 确认这个请求后会给云应用发送一个访问令牌；

（i）云应用向 Google 的服务访问管理模块请求数据，并提交访问令牌；

（j）Google 验证访问令牌有效后，向云应用发送请求的数据。

3. 数据加密技术

云计算中的数据所处的状态包括静态存储状态、传输状态和处理状态。在这三种状态中数据都面临着泄露的风险。针对这三种状态，云服务提供商和研究机构提出了相应的解决方案。

1）数据静态存储加密解决方案

静态数据的加密方案可以使用传统的数据加密方式，但不同的云服务提供商会提供不同的解决方案。亚马逊的简单存储服务并不加密用户数据，但是用户可以在上传数据之前自行加密数据。而微软的云系统则会首先辨别用户数据的敏感度，根据数据敏感度划分数据，将敏感的数据加密。微软的数据加密使用 128 位对称加密方式或 2048 位甚至更长的公开密钥加密机制，所有微软产品必须符合 SDL 加密标准。

2）数据传输加密解决方案

SalesForce 采用 SSL 3.0 或 TLS 1.0 协议保证数据的传输安全，SSL 和 TLS 协议在传输层

对数据进行加密，防止数据被截取和窃听。SSL 和 TLS 协议一般用于用户与云系统之间的安全数据通信，或者是云系统与云系统之间的数据迁移，而云系统内部的安全数据传输方式还没有一个清晰的解决方案。

3）数据处理加密解决方案

因为现有的技术不能有效地直接处理被加密的数据，所以动态数据，即被应用程序直接使用的数据在云中都是未加密的，这就使数据存在泄露的风险。针对这一问题，IBM 的研究员 Gentry 提出了使用理想格（ideal lattices）的完全同态加密算法。完全同态加密机制准许使用者在不解密数据的情况下直接处理加密数据，这种机制使用户的数据在云系统中可以全生命周期处于加密状态，避免了数据泄露的可能。虽然完全同态加密机制在理论上解决了处理加密数据的难题，但是在实际使用中还有许多问题没有解决，需要进一步研究与发展。

4. 入侵检测与 DDoS 防范技术

入侵检测系统（Intrusion Detection System，IDS）与 DDoS 防范技术可以及时发现针对云系统的攻击和云中用户的违规行为，保证云系统的健康运行，它主要解决了云计算的可用性问题，并防止了云计算滥用行为的发生。

入侵检测与行为分析技术就是对系统入侵和违规行为的发现、响应与应对的过程。入侵检测系统通过收集网络和主机中关键点的信息，并对这些信息进行行为分析，从中发现违反安全策略的行为和攻击行为。在云系统中部署入侵检测系统，不仅可以及时发现来自外部的攻击行为，还能针对内部用户进行监控，预防内部用户的违规行为，从而可以解决云资源滥用和云内恶意用户的攻击问题。Sebastian Roschke 等提出了一种在云系统中部署的 IDS，如图 12-6所示。

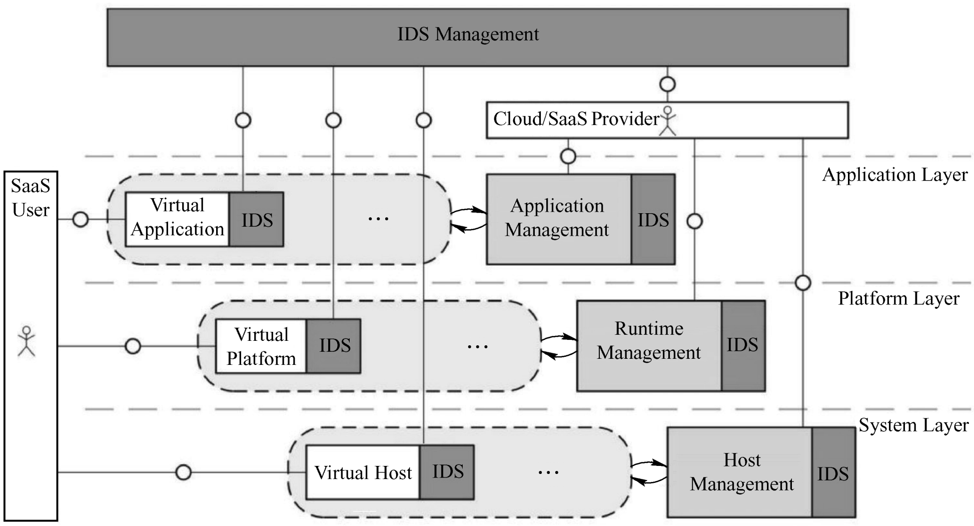


图 12-6　云中入侵检测系统

图 12-6 中的每个虚拟组件都拥有一个 IDS 传感器，这些 IDS 传感器将收集到的数据传送到 IDS 管理组件，管理组件随后分析收集到的数据，当发现攻击行为或违规行为后则采取

对应措施。这些传感器应是用户可见的，并可根据用户自己的意图配置这些传感器及行为库。除了在每个虚拟组件中放置有 IDS 传感器外，云服务提供商要在每一层放置至少一个 IDS 传感器，如在系统层和平台层分别放置基于网络和基于主机的 IDS 传感器，在应用层放置一个基于网络的 IDS 传感器。云服务提供商通过这些传感器，可以很快检测到针对用户的攻击行为或来自云内部用户的攻击。当检测到这些违规行为后，云服务提供商使用自动处置措施及时阻止这些攻击行为，如减少运行 DDoS 攻击的内部用户的资源，或者直接关掉攻击者的主机。DDoS 是对云服务可用性的主要威胁，除了使用入侵检测技术防范 DDoS 攻击外，各个云服务商也采取了不同的措施防范 DDoS 攻击，如 Amazon 采用了专有的 DDoS 攻击缓解技术，可以有效缓解 DDoS 攻击带来的服务中断问题。除此之外，Amazon 的网络都是冗余设计的，通过冗余的服务器提供多点接入，可以在某一点被攻击后，仍然提供服务接入。SalesFoce 则采用冗余链路和维持高带宽的策略应对 DDoS 攻击及可能发生的网络安全威胁。

5. 虚拟机安全技术

虚拟技术是云计算赖以生存的核心技术，所有的公有云服务提供商都通过虚拟技术向用户提供相应的服务。虚拟技术主要面临两个问题：虚拟机本身的安全和用户虚拟机之间的隔离问题。针对这两个问题，研究者们提出了一些解决方案。在 IaaS 服务模式中，云服务提供商会向用户提供一个虚拟机作为用户软件的部署容器，同时也会要求用户及时更新自己的系统和软件，以防范可能出现的软件漏洞。但是由于用户水平与行为的差异，用户不能及时更新维护自己的软件，会造成在云中大量的虚拟机没有及时打上最新漏洞的补丁，从而直接威胁整个云平台的安全。为了解决这个虚拟机补丁更新问题，文献“Checking running and dormant virtual machines for the necessity of security updates in cloud environments”提出了一种自动检测云中运行和休眠的虚拟机，并统一为虚拟机安全更新的方法。整个系统架构如图 12-7所示。

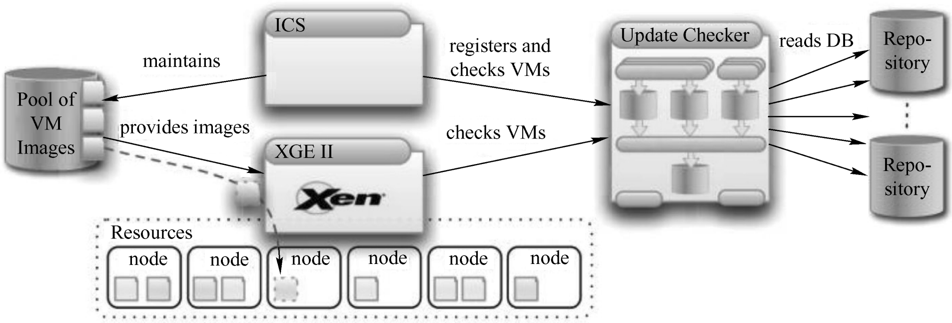


图 12-7　虚拟机软件更新系统架构图

在图 12-7 显示的系统架构中，Update Checker 组件可以有效地检测所有的虚拟机，包括运行中的和休眠的虚拟机，并通过存储在数据库中的软件信息辨别哪些软件需要更新，然后使用软件管理组件更新相应的软件。Update Checker 与已经存在的解决方案（XGE 和 ICS）相结合，可以有效管理虚拟机，如禁止运行没有更新的虚拟机或授权用户自己去管理和更新自己的虚拟机。由于使用了灵活的软件管理组件和缓冲技术，该系统具有更广泛的使用范围和更高的效率。针对虚拟机之间的隔离问题，文献“Resource management for isolation

enhanced cloud services”认为最后一级缓存（Last Level Cache，LLC）共享是阻碍细粒度隔离的一个主要原因，因此提出了两种资源管理方法——缓存层次可感知的核心分配方法和划分缓存的页染色方法，以提供性能和安全的隔离。实验证明，这两种资源管理方法能有效地隔离虚拟机共用的缓存接口，从而实现虚拟机之间的相互隔离。

12.5　基于云计算的物联网信息安全服务体系

在美国国家标准技术研究院 NIST 定义的 SPI 模型中，IaaS 位于底层，提供所有云服务必需的处理、存储的能力；PaaS 建立在 IaaS 之上，为用户提供平台级的服务；SaaS 又以 PaaS 为基础，提供应用级的服务。这 3 种服务模式是功能性、扩展性和复杂性的折中，用户可以根据自身业务特点和需求，选择合适的云服务模式。

从安全角度考虑，云计算安全是云服务提供商和云用户的共同责任，不同的云计算服务模式意味着不同的安全内容和责任划分。云计算安全服务架构的一个关键特点是云服务提供商所在的层次越低，云用户自己所要承担的安全管理职责就越多，如图 12-8 所示。

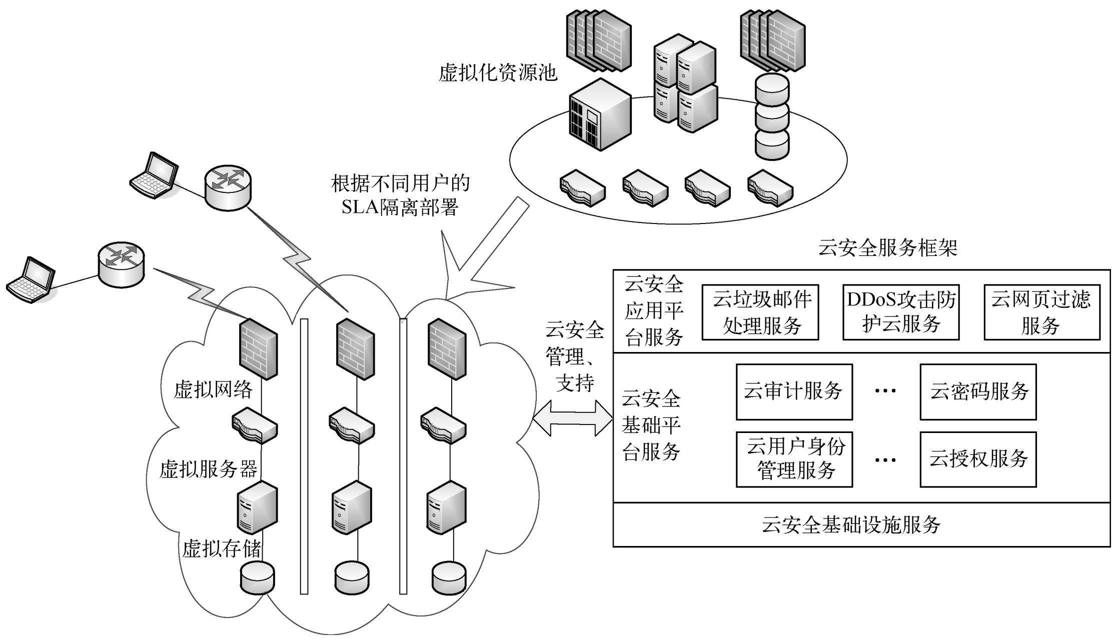


图 12-8　云安全服务框架

在 SaaS 模式下，用户一般只实现云终端的安全，而在服务级别协议（SLA）中对服务提供商的服务水平、安全、管理、合规性等方面提出了明确的要求；但在 PaaS 或 IaaS 模式下，提供商一般仅负责实现底层平台和基础设施组件的功能性和安全性，更多的系统平台和应用层的安全管理责任由用户承担。

云计算是一种计算资源共享架构，云服务提供商统一管理网络设备、服务器、存储设备、安全设备等大量的物理资源，并将这些资源虚拟化，形成一个巨大的资源池。云是一些分布式的并行系统，这些系统由一系列基于服务级别协议（SLA）部署的虚拟计算组成。云计算服务是指用户终端通过远程连接到虚拟机，获取存储、计算、数据库等计算资源或应用

服务。云计算安全服务则是云提供商为实现云用户安全目标（如数据安全、隐私保护和安全管理等），为用户提供的信息安全手段、支持手段，其本质上也是一种云计算服务。对应于云计算服务 SPI 模型，云计算安全服务体系由云计算安全基础设施服务、云计算安全基础平台服务、云计算安全应用服务 3 个安全服务层次构成。

参考文献

[1] FOSTER Ian, ZHAO Yong, RAICU I, etal. Cloud computing and grid computing 360 – degree compared[C]. Grid Computing Environments Workshop, 2008. GCE 08, 2008: 1 – 10.

[2] VAQUERO L M, RODERO – MERINO L, CACERES J, etal. A break in the clouds: towards a cloud definition [J]. ACM SIGCOMM Computer Communication Review, 2009, 39 (1): 50 – 55.

[3] 李爽. 基于云计算的物联网技术研究[D]. 合肥：安徽大学，2014.

[4] 徐小涛，杨志红，等. 物联网信息安全[M]. 北京：人民邮电出版社，2012.

[5] 邓谦. 基于 Hadoop 的云计算安全机制研究[D]. 南京：南京邮电大学，2013.

[6] SCHWARZKOPF R, SCHMIDT M, STRACK C, etal. Checking running and dormant virtual machines for the necessity of security updates in cloud environments[C]. Proceedings – 2011 3rd IEEE International Conference on Cloud Computing Technology and Science, CloudCom 2011, 2011: 239 – 246.

[7] RAJ H, NATHUJI R, SINGHA, etal. Resource management for isolation enhanced cloud services[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, CCSW' 09, Co – located with the 16th ACM Computer and Communications Security Conference, CCS' 09, 2009: 77 – 84.

[8] [34] ALMULLA S A, YEUN C Y. Cloud computing security management[C]. 2010 2nd International Conferenceon Engineering System Management and Applications, ICESMA 2010, 2010: 1 – 7.

[9] NORDBOTTEN N. XML and web services security standards[J]. IEEE Communications Surveys and Tutorials, 2009, 11 (3): 22 – 36.

[10] SHEN Qingni, ZHANG Lizhe, Yang Xin, etal. Securing data migration between cloud storage systems[C]. Proceedings – IEEE 9th International Conference on Dependable, Autonomic and Secure Computing, DASC2011, 2011: 636 – 641.

[11] GENTRY C. Fully homomorphic encryption using ideal lattices[C]. Proceedings of the Annual ACM Symposium on Theory of Computing, 2009: 169 – 178.

[12] ROSCHKE S, CHENG Feng, MEINEl C. Intrusion Detection in the Cloud[C]. 8th IEEE International Symposium on Dependable, Autonomic and Secure Computing, DASC 2009, 2009: 729 – 734.

[13] Amazon. AWS security whitepaper[EB/OL]. http://awsmedia.s3.amazonaws.com/pdf/AWS_Security_Whitepaper.pdf, 2012 – 12 – 15.

[14] BEDRA A. Getting started with Google App engine and clojure[J]. IEEE Internet Computing, 2010, 14 (4): 85 – 88.

[15] GENS F. IT cloud services user survey, pt. 2: top benefits & challenges[EB/OL]. http://blogs.idc.com/ie/?p=210,2012 – 12 – 01.

[16] 续晓燕，丛雪. 浅谈云计算和物联网的融合发展[J]. 电脑知识与技术，2012，08（24）：5782 – 5784.

[17] 韩燕波，赵卓峰，王桂玲，等. 物联网与云计算[J]. 中国计算机学会通讯，2010，6（2）：58 – 63.

[18] 张为民. 物联网与云计算[M]. 北京：电子工业出版社，2012.